

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 24, 2003, 14:03:28 ; Search time 13 seconds
(without alignments)
3886.013 Million cell updates/sec

Title: US-10-021-323-13
Sequence: 1 ggaatgaatcaactttt.....cgaagctncaaatraagg 609

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool/US1002133/runat_24062003.102234_20466/app.query.fasta_1.775
-DB=SwissProt 40 -OPMT=fasten -SUFFIX=esp -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=Blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPM=pcp -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=US1002133.0CGN 1 1 26 @runat_24062003.102234_20466 -NCPU=6 -ICPU=3
-NO_WAMP -LARGEOUTERY -NEG_SCORES=0 -WAIT -DSFBLCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: SwissProt_40*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	239.5	22.6	205	1	ALB1_BETVE
2	218	20.5	161	1	TCH2_ARATH
3	207.5	19.6	199	1	Q09010 golium tub
4	185.5	17.5	171	1	ALB8_OLEEU
5	175	16.5	149	1	CALM_CANAL
6	159	15.0	149	1	CALM_CANAL
7	157	14.8	148	1	CALM_PLAFA
8	155.5	14.7	159	1	TPCS_RABIT
9	155	14.6	148	1	CALM_TERYP
10	155	14.6	151	1	CALM_PNECA
11	154.5	14.6	159	1	TPCS_MOUSE
12	154	14.5	148	1	CALM_EUGGR
13	153.5	14.5	162	1	TPCS_CHICK
14	152.5	14.4	159	1	TPCS_HUMAN
15	152.5	14.4	162	1	TPCS_PANES
16	152	14.3	148	1	CALM_METS
17	152	14.3	148	1	CALM_PATSP
18	152	14.3	148	1	CALM_PATSP

C 19	151.5	14.3	159	1	TPCS_PIG	P02587 sus scrofa
C 20	151	14.2	148	1	CALM_HUMAN	P27482 homo sapien
C 21	151	14.2	148	1	CALM_CHICK	P02597 gallus gall
C 22	151	14.2	162	1	CALM_CHICK	P04352 chlamydomon
C 23	150	14.1	193	1	TPCS_DROME	P48593 drosophila
C 24	149.5	14.1	162	1	TPCS_MELGA	P10246 melagris g
C 25	149	14.0	173	1	CALM_BLAEM	09h1y6 blaetocisti
C 26	148.5	14.0	148	1	CALM_BOVIN	09n1g8 bos taurus
C 27	148	13.9	148	1	CALM_PLEOS	09n1g9 pleurotus o
C 28	148	13.9	165	1	CALM_PLEOS	09n1g9 bos taurus
C 29	147.5	13.9	165	1	POC2_JUNOX	064943 juniperus o
C 30	147	13.9	173	1	CALM_HUMAN	09n986 homo sapien
C 31	147	13.9	148	1	CALM_PLEOS	P27161 phytophthor
C 32	147	13.9	148	1	CALM_PLEOS	P11120 pleurotus c
C 33	147	13.9	148	1	CALM_PLEOS	P11121 pyruidae sp
C 34	146.5	13.8	148	1	CALM_PLEOS	P18061 trypanosoma
C 35	146.5	13.8	148	1	CALM_PLEOS	09j1m3 mus musculu
C 36	146.5	13.8	148	1	CALM_PLEOS	P04630 caenorhabdi
C 37	146	13.8	148	1	CALM_PLEOS	P04465 trypanosor
C 38	145	13.7	148	1	CALM_PLEOS	P13868 solanum ti
C 39	145	13.7	215	1	CALM_PLEOS	09j1k4 mus musculu
C 40	144.5	13.6	173	1	CALM_PLEOS	09j1k3 mus musculu
C 41	144.5	13.6	215	1	CALM_PLEOS	P30188 arabidopsis
C 42	144	13.6	149	1	CALM_PLEOS	P27161 lycopersico
C 43	144	13.6	149	1	CALM_PLEOS	P04464 petunia hyb
C 44	144	13.6	183	1	CALM_PLEOS	P27164 petunia hyb
C 45	144	13.6	219	1	CALM_PLEOS	09n9b3 homo sapien

ALIGNMENTS

RESULT 1
ID ALB1_BETVE STANDARD; PRT; 205 AA.
AC P43187;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcium-binding allergen Bet v 3 (Bet v III).
GN BETVIT1.
OS Betula verrucosa (White birch) (Betula pendula).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae;
OC Eucosmids I; Fagales; Betulaceae; Betula.
OX NCBI_TaxId=3505;
RN [1]
RP SEQUENCE FROM N.A., AND SYNTHESIS OF CALCIUM-BINDING SITES.
RC MEDLINE=94341260; PubMed=7520389;
RX Selberler S., Scheiner O., Kraft D., Lonsdale D., Valenta R.;
RA "Characterization of a birch pollen allergen, Bet v III, representing
RT a novel class of Ca2+ binding proteins: specific expression in mature
RT pollen and dependence of patients' IgE binding on protein-bound
RT Ca2+";
RL EMBO J. 13:3481-3486(1994).
CC -!- FUNCTION: COULD BE INVOLVED IN CALCIUM METABOLISM IN POLLEN. BINDS
CC 3 CALCIUM IONS.
CC -!- SIMILARITY: CONTRAINS 3 EF-HAND CALCIUM-BINDING DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: X29627; CA55854.1; -
CC HSSP; P02593; ICDM.
CC InterPro; IPR002048; EF-hand.
CC Pfam; PF00036; ehand; 3.
CC ProDom; PD000012; EF-hand; 2.
CC SMART; SM00054; Efh; 3.

Author: Selberler et al

RC STRAIN=cv, Columbia;
RA Johnson K.A., Braam J.

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 24, 2003, 13:53:42 ; Search time 38 Seconds

(without alignments) 4271.033 Million cell updates/sec

Title: US-10-021-323-13

Perfect score: 1100
Sequence: 1 ggaatgaacaaacttctt.....tgaagcttacaataaag 609

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ n2p.model -DEV=x1p
-O=/cn2.1/USPTO.spool/US10021322/runat.24062003.102233.20434/app.query.fasta.1.775
-DB=A.GeneSeq.101002 -OPMT=fastan -SUFFIX=rag -MTMARCH=0.1 -LOPCU=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=humanr40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10021322 @CN 1.114 @runat.24062003.102233.20434 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -NARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A.GeneSeq.101002:.*
1: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:.*
2: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:.*
3: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:.*
4: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:.*
5: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:.*
6: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:.*
7: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:.*
8: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:.*
9: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:.*
10: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:.*
11: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:.*
12: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:.*
13: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:.*
14: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:.*
15: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:.*
16: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:.*
17: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:.*
18: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:.*
19: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:.*
20: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:.*
21: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:.*
22: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:.*
23: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	241.5	22.8	169	21	AA07957	Arabidopsis thalia
2	241.5	22.8	169	21	AA043188	Arabidopsis thalia
3	241.5	22.8	215	21	AA07956	Arabidopsis thalia
4	241.5	22.8	215	21	AA043187	Arabidopsis thalia
5	239	22.5	206	20	AA125651	Betula sp. alleege
6	238.5	22.4	209	21	AA023996	Arabidopsis thalia
7	237.5	22.5	154	21	AA023997	Arabidopsis thalia
8	218	20.5	161	21	AA037033	Arabidopsis thalia
9	218	20.5	161	21	AA048187	Arabidopsis thalia
10	218	20.5	161	21	AA048197	Arabidopsis thalia
11	218	20.5	171	21	AA048196	Arabidopsis thalia
12	216.5	20.4	144	21	AA07958	Arabidopsis thalia
13	216.5	20.4	144	21	AA043189	Arabidopsis thalia
14	216	20.4	147	21	AA037034	Arabidopsis thalia
15	216	20.4	147	21	AA048188	Arabidopsis thalia
16	216	20.4	147	21	AA048198	Arabidopsis thalia
17	208.5	19.7	143	21	AA023998	Arabidopsis thalia
18	194	18.3	192	22	AA063709	N tabacum rgs-CaW.
19	193	18.2	157	21	AA004271	Arabidopsis thalia
20	193	18.2	171	21	AA004270	Arabidopsis thalia
21	192	18.1	163	21	AA004327	Arabidopsis thalia
22	192	18.1	167	21	AA004326	Arabidopsis thalia
23	191	18.0	145	21	AA004272	Arabidopsis thalia
24	187	17.6	145	21	AA004328	Arabidopsis thalia
25	163.5	15.4	185	21	AA048314	Arabidopsis thalia
26	163.5	15.4	194	21	AA048313	Arabidopsis thalia
27	162.5	15.3	185	21	AA004158	Arabidopsis thalia
28	162.5	15.3	195	21	AA004157	Arabidopsis thalia
29	155	14.6	150	21	AA095249	Soybean calmodulin
30	154.5	14.6	105	21	AA037035	Arabidopsis thalia
31	154.5	14.6	105	21	AA048189	Arabidopsis thalia
32	154	14.5	150	21	AA021078	Arabidopsis thalia
33	154	14.5	159	23	AA083313	Tropomun C. Unde
34	154	14.5	164	21	AA021077	Arabidopsis thalia
35	152.5	14.4	160	18	AA022597	Human fast twitch
36	152.5	14.4	160	21	AA000134	Human fast twitch
37	152	14.3	138	21	AA021079	Arabidopsis thalia
38	151	14.2	120	21	AA015533	Arabidopsis thalia
39	151	14.2	181	21	AA015532	Arabidopsis thalia
40	150	14.1	166	21	AA013319	Arabidopsis thalia
41	150	14.1	191	21	AA013318	Arabidopsis thalia
42	149.5	14.1	209	21	AA053421	Arabidopsis thalia
43	149.5	14.1	212	21	AA053420	Arabidopsis thalia
44	148	13.9	151	9	AA080162	Biochemical multi
45	147.5	13.9	177	21	AA008198	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AA07957
ID AA07957 standard; Protein; 169 AA.

XX AA07957;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 5306.

XX Protein identification; signal transduction pathway; metabolic pathway;

XX hybridisation assay; genetic mapping; gene expression control; promoter;

XX termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.


```

PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Alignment Scores:

```

Pred. No.: 6.56e-18
Score: 241.50
Percent Similarity: 53.42%
Best Local Similarity: 35.40%
Query Match: 22.76%
DB: 21
Gaps: 4

```

US-10-021-323-13 (1-609) x AAG07957 (1-169)

```

538 CTTAGTGAAGACCGACTTGAACCGGATTCGAGAAGCTCGACAAGATGAGATGCTTC 479
DB 15 ILeaSPProSeRserIleuLeuYsaRgValPheGluMeCpheSPlysaNgIYAspGlyArg 34
QY 478 GTTACGCTGGAGAGAGCTGAATTCGTTGCTCCAGAGAATCGGCTGTCTCAATTCAGCCTT 419
DB 35 ILeThLySgluNdIluLeuNaSPseRLeuGluAsnLeuY--IleTyRileProasp 53
QY 418 GAAGAATTGAGAGCCCTTAGTG-----GGAAACATGTTGAAGCTG 377
DB 54 LysAsPleuThRinIleHsIlySIIeAsPaIAsnIlyAspGlyCySValAspIle 73
QY 376 GATGAATTCGTTGCTCTTATGAATTCATCTCGAACCACTGACATGCGTGTGACGAA 317
DB 74 AspGluPheGluSerIeuTyRseRserIleVal-----AspGlu 86
QY 316 GAGAGAGAGAGAAATTCGATTCACGGCGCGGTGAAGAGAAGACAGTACCTTCCG 257
DB 87 HIs-----HIsAsnAspGlyGluThRgluGluAspMetIys 99
QY 256 AAGGCTTTTAAAGTGTTCGATTCGATGGGAGATGGGAGATGAGAGCTTGA 197
DB 100 AsPaIaPheAsnValPheAspGluAspGlyAspGlyPheIleThRValGluGluLeuYs 119
QY 196 TACGTCGCTGGAGAAAGACTGGGTATGTAGCGTGAATAATGTGAAAAGACTGCAGAGCATG 137

```

```

DB 120 SerValMeCAlaSerIeuIlySgInGlyIySthIleuAspGlyCySblySlyMet 139
QY 136 ATTTGGTATTACGACCAATTCAGACGGCATGCTGATTTTCAGAAATTCAAAAATG 77
DB 140 ILeMeGluValAspIaAspIaAspGlyAspGlyArgValAsnTyRlySgluPheIleuGluMet 159
QY 76 ATG 74
DB 160 Met 160

RESULT 2
ID AAG43188 standard; Protein, 169 AA.
XX
AC AAG43188;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 53951.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KN termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.

```

PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139753.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139839.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 28-JUN-1999; 99US-0140981.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0146389.
 PR 04-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 05-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160711.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
 Pred. No.: 6.56e-18
 Score: 241.50
 Percent Similarity: 53.42%
 Best Local Similarity: 35.40%

Length: 169
 Matches: 57
 Conservative: 29
 Mismatches: 54

Query Match: 22.76% Indels: 21
DB: 21 Gaps: 4
US-10-021-323-13 (1-609) x AAC43188 (1-169)

QY 538 CTTAGTAAAGCCGACTTGGACGCGGTATTGAGAACCTCGACAAAGATGAGATGCTTC 479
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB 15 ILeaApProSeRGlueuLeuYsAtgValPheGlnMetPheAspYsAnGlyAspGlyArg 34
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 478 GTTAGCTGAGAGAGCTGAATGTTGTTCTCCAGAGATGGGTTCTCCATTCAGCCTT 419
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB 35 ILeTrrLySGluGluLeuAnAspSerLeuGluAsnLeuGly---IleTyrrIleProAsp 53
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 418 GAAGATTCGAGCCCTAGT-----GAAACCATGTTGAACCTG 377
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB 54 LysAspLeuThGlnMetIleHsLysIleAspAlaAsnGlyAspGlyCysValAspIle 73
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 376 GATGAATTCCTGTTCTTTATGATTCATCTCGAACCCACTGGACATGCTGTGACGAA 317
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB 74 AspGluPheGluSerLeuYrSerSerIleVal-----AspGlu 86
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 316 GAGAGAGAGAGAAATGTCATTTCACGGCGCGGTGAAGAGAACAGATGACCTTCCG 257
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB 87 HIs-----HIsAsnAspGlyGluThrGluGluGluAspMetLys 99
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 256 AAGGCTTTAAAGTGTTCATTCGATGGGATGGGATGGGATGGGATGGGATGGGATGAA 197
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB 100 AspAlaPheAsnValPheAspGlnAspGlyAspGlyPheIleThrValGluGluLeuLys 119
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 196 TACGTCTGGAGAAAGCTGGATGTGAGGTGAATAATAGTGAAGAAAGATCGACAGACATG 137
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB 120 SerValMetAlaSerLeuGlyLeuYsgInGlyLysThrLeuAspGlyCysIleLysMet 139
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 136 ATTTGATTAACGACCAATTCAGACGGCATGCTGTTTCAAGAAATTCAAAAATG 77
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB 140 IleMetGlnValAspAlaAspGlyArgValAsnTyrrLysGluPheLeuGlnMet 159
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 76 ATG 74
|||
DB 160 Met 160

RESULT 3
AAG07956
ID AAG07956 Standard; Protein; 215 AA.
AC AAG07956;
XX 17-OCT-2000 (first entry)

u2 Arabidopsis thaliana protein fragment SEQ ID NO: 5305.
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridization assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.

```

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145142.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151103.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154479.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

```

```

PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 7.14e-18
Score: 241.50
Percent Similarity: 53.42%
Best Local Similarity: 35.40%
Query Match: 22.76%
DB: 21
Gaps: 4

US-10-021-323-13 (1-609) x AAG07956 (1-215)

QY 538 CTTAGTAAGACCGACTTGCAACCGCTATTGCAAGACTTGACAGAAATGAGATGCTTC 479
DB 61 lEasprrosergluenuhsarvalPhegimmetPheAspLysasnclYAspGlyArg 80
QY 478 GTTAGCTGAGAGAGATGATGTTGCTCCAGAGATGGGCTGCTCAATTCAGACCTT 419
DB 81 lIethrLygluInuhsnspserLeuGluasnleuclY---lIethrIleProasp 99
QY 418 GAAGATTGAGACCCCTTAGTG-----GGAAAACCATGTTGAACCTG 377
DB 100 LysAspLeuThrcIimetiLeHsLysIleAspAlaasnclYAspGlyCysValaspIle 119
QY 376 GATGAATTCCTGTTCTTTATGAAATCCATCTGAAACCCACTGGACATGCTGTGACGAA 317
DB 120 AspGluPhegluSerLeuYrserSerIleVal-----AspGlu 132
QY 316 GAGAGAGAGAGAAATGCTCATTCACGGCGCGGAGAGAAAGACAGACAGTACCTGGG 257
DB 133 HIs-----HIsAsnAspGlyGluThrcIuInuaspmetLys 145
QY 256 AAGGCTTTAAAGTGTGTTGACTTGAATGGAGATGGCTGGCGGAGTGTGAGAGAGCTGAA 197
DB 146 AspAlaPheAsnValPheAspGluAspGlyAspGlyPheIleThrcValGluInuLeuLys 165
QY 196 TACGTGCTGGAAGAGACTGGGTATGTGAGGTGAAATAGTGAAGAAAGACTGCAGAGCATG 137
DB 166 SerValMetAlaSerLeuclYLeuYsgInglYlYsThrcLeuAspGlyCysLysLysMet 185
QY 136 ATTTGTTATTAGACACCAATTCAGACGGCAGTGTGATTTCCAGAAATTCAGAAACATG 77
DB 186 lIemecGlnValAspAlaAspGlyArgValAsnTrpLysGluInuLeuInuMet 205
QY 76 ATG 74
DB 206 Met 206

RESULT 4
AAG43187

```

PR	18-UTN-1999;	99US-0139461;
PR	18-UTN-1999;	99US-0139462;
PR	18-UTN-1999;	99US-0139463;
PR	18-UTN-1999;	99US-0139750;
PR	18-UTN-1999;	99US-0139763;
PR	21-UTN-1999;	99US-0139817;
PR	22-UTN-1999;	99US-0139899;
PR	23-UTN-1999;	99US-0140353;
PR	23-UTN-1999;	99US-0140354;
PR	24-UTN-1999;	99US-0140695;
PR	28-UTN-1999;	99US-0140823;
PR	29-UTN-1999;	99US-0140991;
PR	30-UTN-1999;	99US-0141287;
PR	01-UTL-1999;	99US-0141842;
PR	01-UTL-1999;	99US-0141854;
PR	02-UTL-1999;	99US-0142055;
PR	06-UTL-1999;	99US-0142390;
PR	08-UTL-1999;	99US-0142820;
PR	08-UTL-1999;	99US-0142820;
PR	12-UTL-1999;	99US-0142877;
PR	13-UTL-1999;	99US-0143352;
PR	14-UTL-1999;	99US-0143524;
PR	15-UTL-1999;	99US-0144005;
PR	16-UTL-1999;	99US-0144085;
PR	16-UTL-1999;	99US-0144086;
PR	19-UTL-1999;	99US-0144325;
PR	19-UTL-1999;	99US-0144325;
PR	19-UTL-1999;	99US-0144332;
PR	19-UTL-1999;	99US-0144332;
PR	19-UTL-1999;	99US-0144332;
PR	19-UTL-1999;	99US-0144334;
PR	19-UTL-1999;	99US-0144352;
PR	20-UTL-1999;	99US-0144352;
PR	20-UTL-1999;	99US-0144682;
PR	20-UTL-1999;	99US-0144684;
PR	20-UTL-1999;	99US-0144684;
PR	21-UTL-1999;	99US-0144814;
PR	21-UTL-1999;	99US-0145086;
PR	21-UTL-1999;	99US-0145088;
PR	22-UTL-1999;	99US-0145085;
PR	22-UTL-1999;	99US-0145087;
PR	22-UTL-1999;	99US-0145089;
PR	22-UTL-1999;	99US-0145192;
PR	22-UTL-1999;	99US-0145192;
PR	23-UTL-1999;	99US-0145218;
PR	23-UTL-1999;	99US-0145224;
PR	26-UTL-1999;	99US-0145576;
PR	27-UTL-1999;	99US-0145513;
PR	27-UTL-1999;	99US-0145518;
PR	27-UTL-1999;	99US-0145519;
PR	28-UTL-1999;	99US-0145591;
PR	02-AUG-1999;	99US-0146386;
PR	02-AUG-1999;	99US-0146388;
PR	02-AUG-1999;	99US-0146589;
PR	02-AUG-1999;	99US-0146589;
PR	04-AUG-1999;	99US-0147038;
PR	04-AUG-1999;	99US-0147038;
PR	04-AUG-1999;	99US-0147038;
PR	04-AUG-1999;	99US-0147038;
PR	05-AUG-1999;	99US-0147192;
PR	05-AUG-1999;	99US-0147203;
PR	06-AUG-1999;	99US-0147760;
PR	06-AUG-1999;	99US-0147716;
PR	09-AUG-1999;	99US-0147933;
PR	09-AUG-1999;	99US-0147935;
PR	10-AUG-1999;	99US-0148171;
PR	10-AUG-1999;	99US-0148319;
PR	12-AUG-1999;	99US-0148341;
PR	13-AUG-1999;	99US-0148365;
PR	13-AUG-1999;	99US-0148684;
PR	16-AUG-1999;	99US-0149368;
PR	17-AUG-1999;	99US-0149175;
PR	18-AUG-1999;	99US-0149626;
PR	20-AUG-1999;	99US-0149722;
PR	20-AUG-1999;	99US-0149722;
PR	20-AUG-1999;	99US-0149929;
PR	23-AUG-1999;	99US-0149902;

```

PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158232.
PR 13-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 7.14e-18
Score: 241.50
Percent Similarity: 53.42%
Best Local Similarity: 35.40%
Query Match: 22.76%
DB: 21
Gaps: 4
US-10-021-323-13 (1-609) x AAG3187 (1-215)
QY 538 CTTAGTAAGCCGACTTTCGCAACGCTTTCGAGAAGCTCGACAAGATGAGATGCTTC 479
DB 61 IIAAPProSerIuLeuLysArgValPheGlnMetCpheaSpLysAnGlyAspGlyArg 80
QY 478 GTTAGTCTGAGAGAGCTGATTTGCTCTCCAGAGAAGTGGGCTGTCCATTCAGGCTT 419
DB 81 IIAethLysGlnLuuLeuAenapSerLeuGlnLuuLeuLys---IIeTyrlleProAsp 99

```

```

QY 418 GAAGAATTGAGCCTTACTG-----GAAAAACCATGTTGAACCTG 377
DB 100 LysAspLeuThrGlnMetIleHisLysIleAspAlaAsnGlyAspGlyCysValAspIle 119
QY 376 GATGAATTCCTGTTCTTTATGAATTCATCTGCAACCCACTGGCAATGGTGTAGCGAA 317
DB 120 AspGlnPheGlnSerLeuTyrsSerIleVal-----AspGlu 132
QY 316 GAGAGAGAGAGAAATTTGTCATTCACGGCGCGGTGAAGAAGACAGTGAACCTTGGC 257
DB 133 His-----HisAnaAspGlyGlnThrGlnLuuLuuAspMetLys 145
QY 256 AAGCTTTAAAGTTTGAATTCGTAATGGAGATGGGTGGGGAGTGTAGAGACTTGA 197
DB 146 AspAlaPheAsnValPheAspGlnAspGlyAspGlyPheIleThrValGlnGlnLeuLys 165
QY 196 TACGTCTGGAAGACTGGGTATGTGAGTGAAATAAGTGAAGAAAGACTGCAAGACATG 137
DB 166 SerValMetAlaSerLeuGlnLysGlnGlyLysThrLeuAspGlyCysLysLysMet 185
QY 136 ATTTGGTATTACGACACCAATTCAGAGCGGCATGTGATTTTCAGAAATTCAAAAACATG 77
DB 186 IleMetGlnValAspAlaAspGlyAspGlyArgValAsnTyrlLysGlnLuuLeuGlnMet 205
QY 76 ATG 74
DB 206 Met 206

RESULT 5
AAV25651
AAV25651 standard; protein; 206 AA.
XX
AC AAV25651;
XX
DT 30-SEP-1999 (first entry)
XX
DE Betula sp. allergen 1168696 Bet v 3 protein fragment.
XX
KW Major histocompatibility complex; class II; desensitizing; human;
KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
KW chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity.
XX
OS Betula sp.
XX
PN WO9934826-A1.
XX
PD 15-JUL-1999.
XX
PF 11-JAN-1999; 99WO-GB00080.
XX
PR 21-SEP-1998; 98GB-0020474.
PR 09-JAN-1998; 98GB-0000445.
XX
PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
PI Kay AB, Larche M;
XX
DR WPI, 1999-458255/38.
XX
PT Desensitizing patients to polypeptide allergens
XX
PS Example 6; Page 69; 117pp; English.
XX
CC This invention describes a novel method of desensitizing a patient to a
CC polypeptide allergen and comprises administering to the patient a peptide
CC derived from the allergen where restriction to a MHC class II molecule
CC possessed by the patient can be demonstrated for the peptide and the
CC peptide is able to induce a late phase response in an individual who
CC possesses the MHC class II molecule. The methods can be used for
CC desensitizing patients to allergens present in e.g. grass, tree and weed

```

OS *Arabidopsis thaliana*,

OS *Arabidopsis thaliana*,

[illegible]


```

PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142890.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145818.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.

```

```

PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158025.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Alignment Scores:

```

Pred. No.: 1.52e-17
Score: 238.50
Percent Similarity: 54.768
Best Local Similarity: 33.938
Query Match: 22.488
DB: 21
Gaps: 5

```

US-10-021-323-13 (1-609) x AAC29996 (1-209)

```

QY 547 ATGTCCCTTGAAGACCACTTGCACGCTTTCGAAAGCTGCACAAATGCA 488
   :::::  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 53 LeuLyGlnMeCAsperThrglueuAenATGValPheGlnMeCpheAspLyaspGly 72
QY 487 GATGGCTTTCGTAGCTGTGAGAGCTGATGCTGCTCCAGAAATCGGCTGTGCCA 428
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 73 AspGlyLysIleThrThrglueuAenGluSerPheLysAsnLeuGlyIleIle 91
QY 427 TTCAGCCTTGAAGATTCAGCCCTTATG-----GAAATCCATGT 386
   ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 92 IleProGluAspGluLeuThrglnIleGlnLysIleAspValAsnGlyaspGly 111
QY 385 TTGAAGCTGATGATTCCTTCTTTATGAATCCATTCGAACCCACTGCACATGCT 326
   ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 112 ValAspIleGluGluPheGlyGluLeuTyrlLysIleIleMeC----- 125
QY 325 GGTGACGAGAGAGAGAGGATTCGTCATTCACGGCGGGGTGAAGAGACACT 266
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 126 -----ValGluAspGluAspGluVal-----GlyGluGlu----- 135
QY 265 GACCTTGCAAGGCTTTAAAGTGTTCGCTTGAATGGAGATGGCTGGGGGAGTGTGAG 206

```

Db 136 AspmetylserineValPheAspArgAsnGlyPheIleThrValAsp 155
QY 205 GAGCTTGATACCTGCTGGAGACAGCTGGATGTAGTGAATAAGTGAAGACTGC 146
Db 156 GlutLeuLysLeuValLeuSerSerLeuGlyLeuLysThrLeuGluGluCys 175
QY 145 AGGAGCATGATTGTATTACGACACCAATTACAGCGCATGGTGTATTTCAGAATTTC 86
Db 176 ArgLysMetIleMetGlnValAspValAspGlyAspGlyArgValAsnIleMetGluPhe 195
QY 85 AAAACATGATGTTTACATTCCTCGT 62
Db 196 ArgGlnMetMetLysLysGlyArg 203
RESULT 7
AAG29997
ID AAG29997 standard; Protein; 154 AA.
XX AAG29997;
U 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 35785.
XX
KM Protein identification; signal transduction pathway; metabolic pathway;
KW hybridization assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-014518.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.

PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130649.
PR 23-APR-1999; 99US-0130910.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134841.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136382.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138647.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143524.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-01443352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148341.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.


```
QY 475 AGTCGAGAGAGCTGAATGGTGTCTCCAGAGAACTGGGTCTGTC----- 431
DB 34 SerValAlaSpLeuIleuLysGluValIleArgAlaIleuSerProThrAlaSerProGluGlu 53
QY 430 -----CAATTCAGCCTTGAAGAATTGAGACCCCTTAGTGGGAAACCA 389
DB 54 ThrValThrMetCysGlnPheAspLeuAsp-----GlyAsnGly 67
QY 388 TGTTCGAGCTGATGATGATCTTCTGTTATGATGATCATCTCGAACCCACCTGGACAT 329
DB 68 PheIleAspLeuAspGluPheValAlaLeuPheIle-----IleGlyIle 82
QY 328 GGTGGTGACGAAGAAGAGAGAGAAATTGATTCACGCGCGCGGTGAAGAAGAAC 269
DB 83 GlyGly-----GlyGlyAsnAsnArgAsnAsp 91
QY 268 ---AGTCACCTTGCGAAGCCTTTAAAGTGTGATGATGAGGATGGGTTGGGGGA 212
DB 92 ValSerAspLeuLysGluAlaPheGluLeuTyrAspLeuAspGlyAsnGlyArgIleSer 111
QY 211 TGTGAGAGCTGTAATCGTCTGCGAAGACTGGGTATGTGAGGTGAATAATAGCAAAA 152
DB 112 AlaLysIleuLeuHisSerValMetLysAsnLeuGly-----GluLysCysSerValGln 129
QY 151 GACTGCAGAGACATGATTTGTATTACGACACCAATTGACGCGCATGTTGATTTCAA 92
DB 130 AspCysLysLysMetIleSerLysValAspIleAspGlyAspGlyCysValAsnPheAsp 149
QY 91 GAATTCAAAAACATGATG 74
DB 150 GluPheLysLysMetMet 155

RESULT 10
AAG48197
ID AAG48197 standard; Protein; 161 AA.
XX
AC AAG48197;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 60840.
XX
KM Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
OS Arabidopsis thaliana.
XX
EP1033405-A2.
PU 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 22-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143524.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
```



```

PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145818.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158222.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.

```

```

PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 2,588-15 Length: 161
Score: 218.00 Matches: 55
Percent Similarity: 52.41% Conservative: 32
Best Local Similarity: 33.13% Mismatches: 43
Query Match: 20.55% Indels: 36
Gaps: 6

US-10-021-323-13 (1-609) x AAG48197 (1-161)

QY 535 AGTAAAGCCGACTTTCGACGCGTATTCGAGAACTCGACAGAGATGAGATGCGTTCGTT 476
Db 14 SerMetAspAspLeuLysValPheGlnArgPheAspLysAsnGlyLysIle 33
QY 475 AGCTCGAGAGAGCTGTAATTCGTTCTTCACAGAGATCGGCTGTC----- 431
Db 34 SerValAspGlnLeuLysGlnValIleArgAlaLeuSerProThrLaseProGln 53
QY 430 -----CAATTCAGCCCTGAAGATTGAGAGCCCTAAGGGGAAACCA 389
Db 54 ThrValThrMetLeuLysGlnPheAspLeuAsp-----GlyAsnGly 67
QY 388 TGTTCGACTTCGATGAATTCGTTCTTTATGAATTCATCGAACCCACTGGACAT 329
Db 68 PheIleAspLeuAspIlePheValAlaLeuPheGln-----IleGlyIle 82
QY 328 GGTGTGACCAAGAGAGAGAGAGATTCATTCACGCGCGCGCTGAGAGAAAGAC 269
Db 83 GlyGly-----GlyGlyAsnAsnArgAsnAsp 91
QY 268 ---AGTGACCTTCGCAAGGCTTTAAAGTCTTGAATCGAATGAGGATGGGCGGA 212
Db 92 ValSerAspLeuLysGlnIlePheGlnLeuLysAspLeuAspIleValArgIleSer 111
QY 211 TGTGAGAGCTGAATACGTCGTGGAGAGAGCGGTATGAGGTGAGTGAATACTGAAAA 152
Db 112 AlaLysGlnLeuHisSerValMetLysAsnLeuGly-----GlnLysCysSerValGln 129
QY 151 GACTCGAGAGAGATGATTCGATTACGACACCAATTCAGACGCGATGGTTATTTCAA 92
Db 130 AspCysLysLysMetIleSerLysValAspIleAspGlyAspIlyCysValAsnPheAsp 149
QY 91 GAATTCAAAACATGATG 74
Db 150 GluPheLysLysMetMet 155

RESULT 11
AAG48196 standard; Protein; 171 AA.
XX AAG48196;
AC AAG48196;
XX 18-OCT-2000 (first entry)
DT 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 60839.
DE Protein identification; signal transduction pathway; metabolic pathway;
XX
XX

```

KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134841.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135153.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139465.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144614.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.

PR	25-MAY-1999;	99US-0136021.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138847.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140053.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	20-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145513.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	05-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	06-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	09-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.

PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144336.
 PR 20-JUL-1999; 99US-0144337.
 PR 20-JUL-1999; 99US-0144338.
 PR 21-JUL-1999; 99US-0144339.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.

PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159337.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161921.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
 Pred. No.: 3,63e-15 Length: 144
 Score: 216.50 Matches: 53
 Percent Similarity: 52.03% Conservative: 24
 Best Local Similarity: 35.81% Mismatches: 50
 Query Match: 20.41% Indels: 21
 DB: Gaps: 4

US-10-021-323-13 (1-609) x AAG43189 (1-144)

QY 499 GACAAAGATGAGATGAGCTTGTGAGAGAGCTGAATGAGTGTGCTCCAGAGATC 440
 |||||
 Db 3 AsplysaenGlyAspGlyAArgIlethrIysGlnLeuAsnAspSerLeuGlnAsnLeu 22
 QY 439 GGGTGTGTCATTCAGCCCTTGAAGATTCAGCCCTTAGTG----- 398
 |||
 Db 23 Gly---IleTyxIleProAspIysAspLeuThrGlnMetIleHsIysIleAspIleAsn 41
 QY 397 GGAAGAACATGTTGACCTTGAATTCCTTGTTCCTTTATGATTCATCTCGAACCA 338
 |||||
 Db 42 GlyAspGlyCyAspValAspIleAspGlnPheGlnSerLeuTySerSerIleVal----- 59
 QY 337 CTGGCAGCATGTCGTGTCGAGAGAGAGAGAGATTCGATTCAGCGCGCGGTGAA 278
 |||||
 Db 60 -----AspGlnHis-----HisAsnAspGlyLeu 67
 QY 277 GAAGAGACAGTACCTTGCAGAGCTTTAAAGTGTTCAGCTGAATGAGGATGATG 218
 |||||
 Db 68 ThrGlnGlnGlnAspMetIysAspIleAspValPheAspGlnAspGlyAspGlyPhe 87
 QY 217 GGGGATGTGAGAGCTTGAATTCAGCTGCTGGAAGACCTGGTATGAGGTGAAATAGT 158
 |||||
 Db 88 IleThrValGlnGlnLeuIleAspValMetIleAspLeuGlnGlnGlyLeuThr 107
 QY 157 GGAAGAGATGAGAGATGATTCGATTCAGACCAATTCAGAGCGATGATGAT 98
 |||||
 Db 108 LeuAspGlyCyAspIysMetIleMetGlnValAspIleAspGlyAspGlyValAsn 127
 QY 97 TTTCAGATTCAGAAACATGATG 74
 |||||
 Db 128 TyIysGlnPheLeuGlnMetMet 135
 RESULT 14
 AAG37034

ID AAG37034 standard; Protein; 147 AA.
XX AAG37034;
AC
XX 18-OCT-2000 (first entry)
DT
XX Arabidopsis thaliana protein fragment SEQ ID NO: 45472.
DE
XX Arabidopsis thaliana protein fragment SEQ ID NO: 45472.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridization assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX Arabidopsis thaliana.
OS
XX EP1033405-A2.
PN
XX
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-0301439.
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143524.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147316.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148584.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.


```
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155186.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 4.16e-15
Score: 216.00
Percent Similarity: 52.76%
Best Local Similarity: 33.13%
Query Match: 20.36%
DB: 21
Gaps: 6

US-10-021-323-13 (1-609) x AAG37034 (1-147)
QY 526 GACTTCACAGCGCTTATTCAGAGCTCGACACAAATGAGAGATGCGTTCGTTACTTGAGAG 467
DB 3 AapIleTlyValhegInaRphEaSpLySaSnGlyAaSpGlyLysIleSeRValaSp 22
QY 466 GAGCTGAATGGTGTGCTCCAGAGATCGGGTGTGCTC----- 431
DB 23 GluLeuLysGluValIleatGAlaLeuSeRProThAlaSeRProGluGluTrValTr 42
```

```
QY 430 -----CAATTCACCTTGAAGATTTGAGCCCTTACTGGGAAAACCATGTTGAAC 380
DB 43 MetMetLysGluInhepSeRValaSp-----GlySaSnGlyPheIleAaP 56
QY 379 TTGATGATGATTCCTGTTCTTTATGATCATCTGACACCCACTGGCACATGGTGTGAC 320
DB 57 LeuAaSpGluInheValIaLeuPheGIn-----IleGlyIleGlyGly--- 70
QY 319 GAAGAGAGAGAGAGAGATTTGTCATTACAGCGCGCGGTGAAGAAAGAAC---AGTGAC 263
DB 71 -----GlyGlyAaSaSnAaRghSaSnhepValSeRaP 80
QY 262 CTTCGCAAGGCTTTTAAAGTGTGTTACTGTAATGGGAGATGGGTGGGGAGATGAGAG 203
DB 81 LeuLysGluAlaPheGluLeuTyRhaPLeuAaSpGlyAaSnGlyAaRgIleSeRaLalysGlu 100
QY 202 CTTCGATACGTGCTGGGAGACTGCGGTATGTAGTGAAATAGTGAAATAGCTGACG 143
DB 101 LeuHISerValMetLysSaSnLeuGly-----GluLysCySeRValGInhepCyLys 118
QY 142 AGCATGATTTGGTATTACGACACCAATTCAGACGCGCATGTTGATTTCAAGATTCAAA 83
DB 119 LysMetLISerLysValaSpIleAaSpGlyAaSpGlyCyValaSaSnhepGluInhep 138
QY 82 AACATGATG 74
DB 139 LysMetMet 141

RESULT 15
AAG48188
ID AAG48188 standard; Protein; 147 AA.
XX
AC AAG48188;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 60828.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
```

PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134229.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 26-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 05-AUG-1999; 99US-0147393.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148564.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158359.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 24, 2003, 14:00:42 ; Search time 14.5 Seconds

(without alignments)
2471.525 Million cell updates/sec

Title: US-10-021-323-13

Perfect score: 1100

Sequence: 1 ggtaatagaatcaacttt.....tgaactnccaataaagg 609

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Egapop 6.0 , Egapext 7.0
Delop 6.0 , Delext 7.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ n2p.model -DEV=xip
-Q/cgn2_1/USPTO.spool/US10021323/rnat.24062003.102236.20552/app.query.fasta_1.775
-DB=Issued Patents AA -OPMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOFCU=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdt
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10021323 @CGN 1.1 28 @unat 24062003.102236.20552 -NCPUS=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA.*
1: /cgn2_6/ptodata/1/aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/aa/PTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	159	15.0	149	3	US-08-963-409-3
C 2	155	14.6	150	4	US-09-239-909-2
C 3	152.5	14.4	160	2	US-08-602-941-1
C 4	152.5	14.4	160	3	US-08-961-264-1
C 5	152.5	14.4	160	4	US-09-442-099A-1
C 6	147	13.9	150	4	US-09-239-909-4
C 7	147	13.9	642	2	US-08-818-253-2
C 8	147	13.9	642	4	US-08-818-252-2
C 9	147	13.9	652	2	US-08-818-253-4
C 10	147	13.9	652	4	US-08-818-252-4
C 11	145.5	13.7	456	1	US-08-464-164-2
C 12	145.5	13.7	456	1	US-08-338-057-2

C 13	145.5	13.7	456	2	US-08-668-416-2	Sequence 2, Appl
C 14	145	13.7	149	1	US-08-100-874-2	Sequence 2, Appl
C 15	144.5	13.6	152	3	US-08-963-409-5	Sequence 5, Appl
C 16	144	13.6	642	2	US-08-818-253-6	Sequence 6, Appl
C 17	144	13.6	642	4	US-08-818-252-6	Sequence 6, Appl
C 18	144	13.6	656	2	US-08-818-253-8	Sequence 8, Appl
C 19	144	13.6	656	4	US-08-818-252-8	Sequence 8, Appl
C 20	143	13.5	408	1	US-07-951-715A-21	Sequence 21, Appl
C 21	143	13.5	408	2	US-08-459-448A-21	Sequence 21, Appl
C 22	143	13.5	408	3	US-08-459-595A-21	Sequence 21, Appl
C 23	143	13.5	408	3	US-08-459-504B-21	Sequence 21, Appl
C 24	143	13.5	408	4	US-08-459-444A-21	Sequence 21, Appl
C 25	143	13.5	408	4	US-09-547-422-21	Sequence 21, Appl
C 26	143	13.5	464	1	US-07-951-715A-22	Sequence 22, Appl
C 27	143	13.5	464	2	US-08-459-448A-22	Sequence 22, Appl
C 28	143	13.5	464	3	US-08-459-595A-22	Sequence 22, Appl
C 29	143	13.5	464	3	US-08-459-504B-22	Sequence 22, Appl
C 30	143	13.5	464	3	US-08-459-444A-22	Sequence 0, Appl
C 31	143	13.5	464	4	US-09-547-422-22	Sequence 0, Appl
C 32	141	13.3	149	3	US-08-963-409-4	Sequence 4, Appl
C 33	141	13.3	149	3	US-08-641-873-20	Sequence 20, Appl
C 34	133	12.4	145	4	US-07-951-715A-24	Sequence 24, Appl
C 35	132	12.4	142	1	US-08-459-448A-24	Sequence 24, Appl
C 36	132	12.4	142	2	US-08-459-595A-24	Sequence 24, Appl
C 37	132	12.4	142	3	US-08-459-504B-24	Sequence 24, Appl
C 38	132	12.4	142	3	US-08-459-444A-24	Sequence 24, Appl
C 39	132	12.4	142	4	US-09-547-422-24	Sequence 24, Appl
C 40	132	12.4	142	4	US-08-963-409-1	Sequence 1, Appl
C 41	127.5	12.0	176	3	US-08-963-409-1	Sequence 4, Appl
C 42	121.5	11.5	149	3	US-08-764-563-4	Sequence 4, Appl
C 43	120.5	11.4	390	3	US-08-993-380-4	Sequence 4, Appl
C 44	118	11.1	639	4	US-09-347-801-17	Sequence 17, Appl
C 45	114.5	10.8	177	3	US-08-764-563-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-963-409-3
Sequence 3, Application US/08963409
Patent No. 6046315
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Puri
TITLE OF INVENTION: DISEASE ASSOCIATED CALMODULIN PROTEIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSRO for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/963,409
FILING DATE: Filed Herewith
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0418 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 149 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: GenBank
 CLONE: 385234
 US-08-963-409-3

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	3.88e-10	149	44	39	55	28
Percent Similarity:	159.00					
Best Local Similarity:	50.00%					
Query Match:	26.51%					
	14.99%					

US-10-021-323-13 (1-609) x US-08-963-409-3 (1-149)

```

550 AAATGCCCCCTTAGTAAGACCGACTTCGACCGTATTTCGAGAGCTCGACAGAAAT 491
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4 LyleuThrcgluagluuInleSerGluPheLysGluLalaPheSerLeuPheAsp 23
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 490 GGAAGATGCTTCTTACTGAGAGAGCTGAATGTTGCTCCAGAGATCGG----- 437
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 24 GlysAspGlyThrIleThrIleThylsGluLeuGlyThyValMetArgSerLeuGlyGlnAsn 43
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 436 -----TCTGCCAATTCAGCCTTGAGAGAAATGAGACCCCTTAGTGGGAAAA 389
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 44 ProThrcgluagluuInleAspMetCileAsnGluLeasp---ThraSpGlyAsnGly 62
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 388 TGTTCGATTCGATGATGATCTTGTCTTTATGATTCGATTCGACCACTCGACAT 329
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 63 ThrIleAspPheProGluPheLeuThrcglu-----MetAlaGlySLeuLys--- 78
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 328 GGTGTGTCAGAGAGAGAGAGAGAAATGTCATTCAGCGCGCGGTGAAGAAAGAC 269
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 79 -----AspThrAspThrGluGluGluLeuLeu----- 87
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 268 AGTGAAGCTTCGAGAGCTTTTAAAGTGTTCGATTCGATTCGAGAGCTTCGAGAGAT 209
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 88 -----GluAlaPheArgValPheAspArgAspGlyAspGlyTyrlleSerAla 103
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 208 GAGAGCTTCGATTCGATTCGAGAGAGCTGAGTGTGAGGTGAAATGAGAAAGAC 149
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 104 AspGluLeuArgHisValMetThrAsnLeuGly-----GluLysLeuThrAsnGlu 121
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 148 TGCAGAGCATGATTTGGTATTACAGACCAATTGACAGCGCATGTTGATTTTCAGAA 89
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 122 ValAspGluMetCileArgGluAlaAspIleAspGlyAspGlyGlnIleAsnTyrlleGlu 141
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 88 TTCAAAACATGATGTTA 71
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 142 PheValLysMetMetCile 147
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 2

US-09-239-909-2
 ; Sequence 2, Application US/09239909
 ; Patent No. 6284952
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumbo Petrochemical Co. Ltd.
 ; TITLE OF INVENTION: Transgenic Plants with Divergent SCAM4 or SCAM5 Gene to Achieve M
 ; FILE REFERENCE: P99-2-6
 ; CURRENT FILING DATE: 1999-01-29
 ; EARLIER FILING DATE: 1999-01-08
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: KOPATIN 1.0
 ; SEQ ID NO 2
 ; LENGTH: 150
 ; TYPE: PRT

ORGANISM: G. max calmodulin4 (SCAM4)
 US-09-239-909-2

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	1.13e-09	150	47	29	57	28
Percent Similarity:	155.00					
Best Local Similarity:	47.20%					
Query Match:	29.19%					
	14.61%					

US-10-021-323-13 (1-609) x US-09-239-909-2 (1-150)

```

526 GACTTCGACCGCTATTGAGAGAGCTCGACAGAAATGAGATGCTTCTTACTGAGAG 467
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 12 AspPheLysGluLalaPheGlyLeuPheAspLysAspGlyAspGlyCysIleThrValGlu 31
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 466 GACTGAAATGCTGCTCCAGAGATCGGCTCTGTCATTCAGCCTTGAGAAATGAGAG 407
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 32 GluLeuAlaThrValIleArgSerLeuAspGln---AsnProThrGluGluGluLeuGln 50
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 406 CCTTAGTG-----GGAAAAACATGTTGAACTTGATGATGATCTTG 365
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 51 AspMetCileSerGluValAspAlaAspGlyAsnGlyThrlleGluPheAspGluPheLeu 70
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 364 TTTCTTTATGATTCATTCGACCACTCGACCATGATGATGATGATGATGATGATGAT 305
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 71 -----SerLeuMetAlaLysLysValLysAspThrAspAlaGluGluGlu----- 85
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 304 GAATTCGATTCACCGCGCGGTGAAGAAAGACAGTGCATTCGACCTTCGAAAGCTTTAAA 245
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 86 -----LeuLysGluAlaPheLys 91
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 244 GTGTTGACTGATGATGAGATGAGTGGGAGATGAGAGCTTCGATTCGATTCGATTCGAG 185
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 92 ValPheAspLysAspGlnAsnGlyTyrlleSerAlaSerGluLeuArgHisValMetCile 111
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 184 AGATGGGATGATGAGATGAGTGAAGAAATGAGAAAGCTCGACGAGCATGATTTGGTATTAC 125
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 112 AsnLeuGly-----GluLysLeuThrAspGluGluValGluGluMetCileLysGluAla 129
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 124 GACACCAATTCAGACCGCATGTTGATTTTCAGAAATTCAGAAACATGATGTTACATTC 65
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 130 AspLeuAspGlyAspGlyGlnValAsnTyrlleGluGluPheValLysMetMetCileThrVal 149
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 64 CGT 62
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 150 Arg 150
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 3

US-08-602-941-1
 ; Sequence 1, Application US/08602941
 ; Patent No. 5837680
 ; GENERAL INFORMATION:
 ; APPLICANT: Moses, Marsha A.
 ; APPLICANT: Langer, Robert S.
 ; APPLICANT: Wiedersheim, Dimitri G.
 ; APPLICANT: Wu, Jimin
 ; APPLICANT: Sytkowski, Arthur
 ; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS
 ; TITLE OF INVENTION: COMPRISING TROPONIN SUBUNIT, FRAGMENTS AND ANALOGS
 ; NUMBER OF INVENTIONS: THEREOF AND METHODS OF THEIR USE TO INHIBIT ANGIOGENESIS
 ; NUMBER OF SEQUENCES: 3
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Penite & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: U.S.A.
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible

```

GENERAL INFORMATION:
APPLICANT: Moses, Marsha A.
APPLICANT: Langer, Robert S.
APPLICANT: Wiedersheim, Dimitri G.
APPLICANT: Wu, Imiin
APPLICANT: Sykowski, Arthur
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS
TITLE OF INVENTION: COMPRISING TROPONIN SUBUNITS, FRAGMENTS AND ANALOGS
TITLE OF INVENTION: THEREOF AND METHODS OF THEIR USE TO INHIBIT ANGIOGENESIS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,264
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/602,941
FILING DATE: 16-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,452
REFERENCE/DOCKET NUMBER: 8657-021-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864/9741
INFORMATION FOR SEQ. ID NO.: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 160 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6025331e
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..160
OTHER INFORMATION: /label= Human Fast Twitch Skeletal
OTHER INFORMATION: Muscle Troponin C
US-08-961-264-1

Alignment Scores:
Pred. No.:          Length:      160
Score:              Matches:       43
Percent Similarity: 45.86%         Conservative: 29
Best Local Similarity: 27.39%        Mismatches:   60
Query Match:        Indels:        25
DB:                  Gaps:           4

US-10-021-323-13 (1-609) x US-08-961-264-1 (1-160)

Oy      526 GACTTGCAACGCGTATTCAGAGAAGTCCAAAGATGGAGATGGCTTGTTAGCTGTGAG    467
Db      :::: :::: |||::: :::: |||::: :::: |||::: :::: |||::: :::: |||:::
      19 GluPhelValAlaIalaheAspheAspalaspGlyglYasplIeservValys     38
Oy      466 GAAGCTAATTTGGTGTGCAGAGATCGGGTCTGTGCCAATTCAGACTTGAGAAATTGGAG    407
Db      :||| |::: :::: :::: |||::: :::: |||::: :::: |||::: :::: |||:::
      39 GluLeuGlYThrValMetArgMetLeuGlYInthrPro---ThrlYsgluGluLeuasp     57
Oy      406 CCCTTAGTGTG-----GGAACAACATTTGAACTTGATGATTTCTTG             365
Db      :::: :::: :::: :::: |||::: :::: |||::: :::: |||::: :::: |||:::
      58 AlaAlrIlegluGluValaspGluaspGlysergYlYThrIleaspPhegluGluPheLeu    77
Oy      364 TTCTTTTAGAACCATCTCCGAACCACCTGGAGCATGTGTGTGACGAAGAGAGAGAGAG     305

```

```

Db      78 Valmetmelvalarglnmetlysgluaspalalygilysergluuglu----- 95
Qy      304 GAATGTCATTACAGCGCGCGGTGAAGAAGACAGTGCCTTGCAAGGCTTTAA 245
Db      96 -----Leu1adglucysphearg 101
Qy      244 GTGTTGACTTAATGGGATGGTGGGGGATGTGAGAGCTTGAATACGTCTGGGA 185
Db      102 IlepheaspargasnlaspolyrilleapProgluugluenlaglullephearg 121
Qy      184 AGACTGGTATGTGAGGTGAATAATAGTGAAGAAAGCTGACAGCATGATTGGTTAC 125
Db      122 Alasergly-----GluHivaltlraspgluugluilegluserleuetylsapgly 139
Qy      124 GACACCAATTGACAGCGCATGTTGATTTCAGAAATTCMAAATCATGATG 74
Db      140 Asplysasnashaspolyargilleapheaspglupheleulymetet 156

```

RESULT 5

```

09-442-099A-1
Sequence 1, Application US/09442099A
Patent No. 6465431

```

```

GENERAL INFORMATION:
APPLICANT: Thorm, R.
APPLICANT: Lanzer, M.
APPLICANT: Moses, M.
APPLICANT: Wiedersheim, D.
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS COMPRISING TROPONIN SUBUNITS,
TITLE OF INVENTION: FRAGMENTS AND HOMOLOGS THEREOF AND METHODS OF THEIR USE TO
FILE REFERENCE: 8657-028
CURRENT APPLICATION NUMBER: US/09/442,099A
CURRENT FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: 09/268,274
PRIOR FILING DATE: 1999-03-15
PRIOR APPLICATION NUMBER: 08/961,264
PRIOR FILING DATE: 1997-10-30
PRIOR APPLICATION NUMBER: 08/602,941
PRIOR FILING DATE: 1996-02-16
NUMBER OF SEQ ID NOS: 20
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 160
TYPE: PRF
ORGANISM: Homo sapiens
US-09-442-099A-1

```

Alignment Scores:

```

d. No.: 2,26e-09 Length: 160
Score: 152.50 Matches: 43
Percent Similarity: 45.86% Conservative: 29
Best Local Similarity: 27.39% Mismatches: 60
Query Match: 14.37% Indels: 25
Gaps: 4

```

```

US-10-021-323-13 (1-609) x US-09-442-099A-1 (1-160)

```

```

Qy      526 GACTTCCAAACGGATATTCGAGAAGCTGCAGACAGATGAGATGCTTCCTAGCTCGAG 467
Db      19 Gluphetylalalalalalalalalalalalalalalalalalalalalalalal 38
Qy      466 GAGCTGAATGTGCTCCAGAGATCGGGTGTGTCATTCATTCAGCTTGAAGATTTGAG 407
Db      39 GlubeglyThrValmetmetleuglyGlnthPro---Thrysgluugluuasp 57
Qy      406 CCTTAGTG-----GGAAAAACATGTTGAATTCGATGATGATTTCTTG 365
Db      58 AlalilellegluuValaspgluyseryglThrilleasphhegluugluu 77
Qy      364 TTTCTTTAATGATCCATCTCGAACCACTGGCCATGCTGCTGATCGAAGAGAGAGAG 305
Db      78 Valmetmelvalarglnmetlysgluaspalalygilysergluuglu----- 95

```

```

Qy      304 GAATGTCATTACAGCGCGCGGTGAAGAAGACAGTGCCTTGCAAGGCTTTAA 245
Db      96 -----Leu1adglucysphearg 101
Qy      244 GTGTTGACTTAATGGGATGGTGGGGGATGTGAGAGCTTGAATACGTCTGGGA 185
Db      102 IlepheaspargasnlaspolyrilleapProgluugluenlaglullephearg 121
Qy      184 AGACTGGTATGTGAGGTGAATAATAGTGAAGAAAGCTGACAGCATGATTGGTTAC 125
Db      122 Alasergly-----GluHivaltlraspgluugluilegluserleuetylsapgly 139
Qy      124 GACACCAATTGACAGCGCATGTTGATTTCAGAAATTCMAAATCATGATG 74
Db      140 Asplysasnashaspolyargilleapheaspglupheleulymetet 156

```

RESULT 6

```

US-09-239-909-4
Sequence 4, Application US/09239909
Patent No. 6284952

```

```

GENERAL INFORMATION:
APPLICANT: Kumo Petrochemical Co. Ltd.
TITLE OF INVENTION: Transgenic Plants with Divergent SCAM4 or SCAM5 Gene to Achieve
FILE REFERENCE: P99-2-6
CURRENT APPLICATION NUMBER: US/09/239,909
CURRENT FILING DATE: 1999-01-29
EARLIER APPLICATION NUMBER: EP 99300136.1
EARLIER FILING DATE: 1999-01-08
NUMBER OF SEQ ID NOS: 4
SOFTWARE: KOPATIN 1.0
SEQ ID NO 4
LENGTH: 150
TYPE: PRF
ORGANISM: G. max calmodulins (SCAM5)
US-09-239-909-4

```

Alignment Scores:

```

Pred. No.: 9,54e-09 Length: 150
Score: 147.00 Matches: 46
Percent Similarity: 49.38% Conservative: 33
Best Local Similarity: 28.75% Mismatches: 51
Query Match: 13.85% Indels: 30
Gaps: 6

```

```

US-10-021-323-13 (1-609) x US-09-239-909-4 (1-150)

```

```

Qy      529 ACCGACTTGCAACGGATATTCGAGAAGCTGCAGACAGATGAGATGCTTCCTAGCTCG 470
Db      11 Sergluilegluallalalalalalalalalalalalalalalalalalalalal 30
Qy      469 GAGAGCTGAATGTGCTCCAGAGATCGGGTGTGTCATTCATTCAGCTTGAAGAT 413
Db      31 AspgluuValThrValilleargserleu-----ValGlnasprothrgluuglu 48
Qy      412 TTGAGCCCTTAGTG-----GGAAAAACATGTTGAATTCGATGATGA 371
Db      49 LeuglnaspmetlilasncllualaspalalaspalyanglyThrillegluValglu 68
Qy      370 TTTCTTGTCTTTATGAATTCATCTCGAACCCACTGGACATGCTGTGACCAAGAGAG 311
Db      69 Phelu-----AsnleuMetlalylyMetleugluThrasp 81
Qy      310 GAGAGGAATTTGATTCATTCACGCGCGCGGTGAAGAAGACAGTGCCTTGCAAGCT 251
Db      82 Gluuglu-----Aspleuysgluua 89
Qy      250 TTTAAGCTTTGACTTGAATGGGATGGTGGGGGATGTGAGAGCTTGAATACGTG 191
Db      90 PhelysvalalalaspolyaspolinsnglyrilleaserAlasercluleuarghival 109
Qy      190 CTGGAGAGCTGGGTATGTGAGTGAATAATAGTGAAGAAAGCTGACAGCATGATTTGG 131

```


QY 169 GGTGAATAATGATGGAAGACCTGACAGACATTTGATTAACACACCAATTCAGAC 110
Db 344 GtUlyLeuThraspIuValaAspGluMetIleArgGluAlaAspIleAspGlyAsp 363
QY 109 GGCATGTTGATTTTCAGAAATTCAAAACATGATG 74
Db 364 GlyGlnValaAsnTyrgIuGluPheValGlnMetMet 375

RESULT 9
US-08-818-253-4
Sequence 4, Application US/08818253
Patent No. 5998204
GENERAL INFORMATION:
APPLICANT: Tsien, Roger Y.
APPLICANT: Miyawaki, Atsushi
TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
NUMBER OF INVENTION: DETECTION OF ANALYTES
CORRESPONDENCE ADDRESS: 61
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
FILING DATE: 14-MAR-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Hallé, Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07257/043001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 652 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-818-253-4

Alignment Scores:
Pred. No.: 1.8e-08 Length: 652
Score: 147.00 Matches: 41
Percent Similarity: 50.66% Conservative: 36
Best Local Similarity: 26.97% Mismatches: 57
Query Match: 13.85% Indels: 18
Gaps: 4
DB: 2

US-10-021-323-13 (1-609) x US-08-818-253-4 (1-652)

QY 526 GACTTGCAACGCGTATTTCAGAACTCGACAAAGATGAGCTTGTAGTCTGAG 467
Db 241 GlnPheLysGluAlaPheSerLeuPheAspGlyAspGlyThrIleThrLys 260
QY 466 GAGCTGAATGTTGCTCTCAGAGAAATCGGCTCTGTCATTCAGCCTTGA--GAATTG 410
Db 261 GlnLeuGlyThrValMetArgSerLeuGly-----GlnAsnProthrdGluAlaGluLeu 278
QY 409 GAGCCCTTAGTGGAAGAAACCATGTTGAACCTTGAGATGATCTTGTCTTTATGATCC 350
Db 279 GlnAspMetIleAsnGlnValaAspGlyAsnGlyThrIleTyrrPheProGluPhe 298

QY 349 ATCTGCAACCCACTGCGACATGGTGTGACGAAAGAGAGAGAGAAATGTCATTAC 290
Db 299 LeuThrMetMetAlaArgLysMetLysAspThrAspSerGluGlu----- 314
QY 289 GGGCGGGGGAAGAAACACAGTACCTTGGAAGCGCTTTAAAGTGTGATTCAT 230
Db 315 -----IleArgGluAlaPheArgValPheAspLysAsp 325

QY 229 GGGATGCGTGGGGGATGTGAGAGCTTGAATCGTCTGGAAGACTGGATATGTA 170
Db 326 GlyAsnGlyTyrIleSerAlaAlaGluLeuAlaGlnValMetThrAsnLeuGly----- 343
QY 169 GGTGAATAATGATGGAAGACCTGACAGACATGATTTGATTAACACACCAATTCAGAC 110
Db 344 GtUlyLeuThraspIuValaAspGluMetIleArgGluAlaAspIleAspGlyAsp 363
QY 109 GGCATGTTGATTTTCAGAAATTCAAAACATGATG 74
Db 364 GlyGlnValaAsnTyrgIuGluPheValGlnMetMet 375

RESULT 10
US-08-818-252-4
Sequence 4, Application US/08818252B
Patent No. 6197928
GENERAL INFORMATION:
APPLICANT: Tsien, Roger Y.
APPLICANT: Miyawaki, Atsushi
TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
NUMBER OF INVENTION: DETECTION OF ANALYTES
FILE REFERENCE: 07257/042001
CURRENT APPLICATION DATA:
FILING DATE: 1997-03-14
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 4
LENGTH: 652
TYPE: PRP
ORGANISM: Aequorea victoria
US-08-818-252-4

Alignment Scores:
Pred. No.: 1.8e-08 Length: 652
Score: 147.00 Matches: 41
Percent Similarity: 50.66% Conservative: 36
Best Local Similarity: 26.97% Mismatches: 57
Query Match: 13.85% Indels: 18
Gaps: 4
DB: 4

US-10-021-323-13 (1-609) x US-08-818-252-4 (1-652)

QY 526 GACTTGCAACGCGTATTTCAGAACTCGACAAAGATGAGCTTGTAGTCTGAG 467
Db 241 GlnPheLysGluAlaPheSerLeuPheAspGlyAspGlyThrIleThrLys 260
QY 466 GAGCTGAATGTTGCTCTCAGAGAAATCGGCTCTGTCATTCAGCCTTGA--GAATTG 410
Db 261 GlnLeuGlyThrValMetArgSerLeuGly-----GlnAsnProthrdGluAlaGluLeu 278
QY 409 GAGCCCTTAGTGGAAGAAACCATGTTGAACCTTGATGATCTTGTCTTTATGATCC 350
Db 279 GlnAspMetIleAsnGlnValaAspGlyAsnGlyThrIleTyrrPheProGluPhe 298
QY 349 ATCTGCAACCCACTGCGACATGGTGTGACGAAAGAGAGAGAGAAATGTCATTAC 290
Db 299 LeuThrMetMetAlaArgLysMetLysAspThrAspSerGluGlu----- 314
QY 289 GGGCGGGGGAAGAAACACAGTACCTTGGAAGCGCTTTAAAGTGTGATTCAT 230
Db 315 -----IleArgGluAlaPheArgValPheAspLysAsp 325
QY 229 GGGATGCGTGGGGGATGTGAGAGCTTGAATCGTCTGGAAGACTGGATATGTA 170
Db 326 GlyAsnGlyTyrIleSerAlaAlaGluLeuAlaGlnValMetThrAsnLeuGly----- 343

Db 326 GlyAsnGlyTyrIleSerAlaIaGluLeuArgHisValMetThrAsnLeuGly----- 343
 QY 169 GGTGAATAATAGTGAAGACCTGCAGAGATGATTGGTATTACACACCAATTGAC 110
 Db 344 GluTyrLeuThrAspIleGluValAspGluMetIleArgGluAlaAspIleAspGlyAsp 363
 QY 109 GGCATGTTGATTTCAGAAATTCAAAACATGATG 74
 Db 364 GlyGlnValAsnTyrGluGlnPheValGlnMetMet 375
 RESULT 11
 US-08-464-164-2
 ; Sequence 2, Application US/08464164
 ; Patent No. 5614195
 ; GENERAL INFORMATION:
 ; APPLICANT: Tomley, Fiona M.
 ; APPLICANT: Dunn, Paul P. J.
 ; APPLICANT: Bumstead, Janene M.
 ; APPLICANT: Vermeulen, Arno N.
 ; TITLE OF INVENTION: Coccidiosis poultry vaccine
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Akzo No. 5614195el Patent Department
 ; STREET: 1300 Piccard Drive, Suite 206
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: U.S.A.
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/464,164
 ; FILING DATE: June 2, 1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Gormley, Mary E.
 ; REGISTRATION NUMBER: 34,409
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 258-5200
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 456 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; 08-464-164-2
 Alignment Scores:
 Pred. No.: 2,3e-08 Length: 456
 Score: 145.50 Matches: 51
 Percent Similarity: 41.18% Conservatve: 26
 Best Local Similarity: 27.27% Mismatches: 59
 Query Match: 13.71% Indels: 51
 DB: 1 Gaps: 6
 US-10-021-323-13 (1-609) x US-08-464-164-2 (1-456)
 QY 565 ATCCCTTTGGTGTAAATGTCCTCCCTTAGTAAGACGAC--TTGCAACGGGTATTC 509
 Db 290 LeuTyrMetGlySerIlySleuThrThAsnGluGluThrAspGluLeuAsnIlySlePhe 309
 QY 508 GAGAAGCTGCACAGAAATGAGATGCGTTCGTTGCTCGAGAGAGCTGAATGTTGCTC 449
 Db 310 GlnIlyMetAspIlySaSnGlyAspGlyGlnLeuAspIlySnglnIleu----- 325
 QY 448 CAGAGATCGGGTGTGCATTCAGCCTTGAAGAAATGAGCCCTTAGTGGAAGAAACA 389
 Db 326 -----MetGluGlyTyrValGlnIleuMetIlySleuIlySgIlyGluAsp 339
 QY 388 TGTTTGAACCTTGATGAATCTTGTCTTTATGAATCAATCTGAAACCACTGGACAT 329

Db 340 ValSerValLeuAspIlySerAlaIle----- 348
 QY 328 GGTGTGACGAAGAGAGAGAGAAATGTTGATTCACGGC-----GGC 284
 Db 349 -----GluThrGlnValGluGlnValLeuGluAlaValAspPheAspIlySaSn 364
 QY 283 GGTGAAGAAGAGACAGTGC----- 263
 Db 365 GlyPheIleGluTyrSerGluPheValThrValAlaMetAspArgArgThrLeuLeuSer 384
 QY 262 -----CTTGCAGACTTTTAAAGTGTGACTTGAATGAGGATGGTGGGGGA 212
 Db 385 ArgGlnArgLeuGlnArgAlaPheGluMetPheAspSerAspIlySerGlyIleSer 404
 QY 211 TGTGAGAGCTTGAATACGTCCTGCGAGACCTGGATGTGAGTGAATAATGGAATA 152
 Db 405 SerSerGluLeuAlaThrIlePheGlyVal-----SerGluLeuAspSerGlu 420
 QY 151 GACTGACAGACGTATGTTGTTACGACACCAATTGACAGCGCATGTTGATTTTCAA 92
 Db 421 AlaTrpArgArgValLeuAlaGluValAspArgAsnAspIlyGlnValAspPheGlu 440
 QY 91 GAATCAAAAACATGATGTTA 71
 Db 441 GluPheGlnGlnMetLeuLeu 447
 RESULT 12
 US-08-338-057-2
 ; Sequence 2, Application US/08338057
 ; Patent No. 5795741
 ; GENERAL INFORMATION:
 ; APPLICANT: Tomley, Fiona M.
 ; APPLICANT: Dunn, Paul P. J.
 ; APPLICANT: Bumstead, Janene M.
 ; APPLICANT: Vermeulen, Arno N.
 ; TITLE OF INVENTION: Coccidiosis poultry vaccine
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Organon Teknika Corporation
 ; STREET: 1330 Piccard Drive
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: U.S.A.
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/338,057
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: EP 93.309078.9
 ; FILING DATE: 12-NOV-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Gormley, Mary E.
 ; REGISTRATION NUMBER: 34,409
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 258-5200
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 456 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-338-057-2
 Alignment Scores:
 Pred. No.: 2,3e-08 Length: 456
 Score: 145.50 Matches: 51

Percent Similarity:	41.18%	Conservative:	26
Best Local Similarity:	27.27%	Mismatches:	5
Query Match:	13.71%	Indels:	51
DB:	1	Gaps:	6

US-10-021-323-13 (1-609) X US-08-338-057-2 (1-456)

```

Oy 565 ATCCCTTTGGGGGAAAATATGTCGCCCCCTTAGTAAACCGAGC---TTGGCAACCGCATTC 509
Db 290 LeuTrpMetGlySerLysLeuThrThrAsnGluGluThrAspGluLeuAsnLysIlePhe 309
Oy 508 GAGAAAGCTCGACACAGAAATGAGATGCGTTCGTTAGTCTGGAAGAGCTGAATTGGTGTCTC 449
Db 310 GlnLysMetAspLysAlaAsnGlyAspGlyGlnLeuAspLysGlnGlnLeu----- 325
Oy 448 CAGAGAAATCGGGTCTGTCCAAATTCACGCTTGAAGAAATTGGAGCCCTTAGTGGGAAAACCA 389
Db 326 -----MetGlnGlyTrpValGlnLeuMetLysLeuLysGlyGlnAsp 339
Oy 388 TGTTTGAACCTTGATGATATTCCTTGCTCTTTATGAATCATCATCGAAACCACTGCACAT 329
Db 340 ValSerValLeuAspLysSerAlaIle----- 348
Oy 328 GGTGTGACGAAAGAGAGAGAGAGAAATGTCATTACAGCGC-----GGC 284
Db 349 -----GlnThrGlnValGlnGlnValLeuGlnAlaValAspPheAspLysAsn 364
Oy 283 GGTGAAGAAGAGACAGTAC----- 253
Db 365 GlyPheIleGlnTrpSerLupPheValIThrValAlaMetAspArgTrpThrLeuLeuSer 384
Oy 262 -----CTTGGCGAGCTTTAAAGTGTAGCTTGACCTGAAATGGGATGGGCGGGA 212
Db 385 ArgGlnArgLeuGlnArgAlaPheGlnMetPheAspSerAspLysSerGlyLysIleSer 404
Oy 211 TGTGAGAGAGCTTGAATACCTGCTGGCGAAGACTGGGTATGTGAGGTGAATAAGTGAAAA 152
Db 405 SerSerGlnLeuAlaThrIlePheGlyVal-----SerGlnLeuAspSerGln 420
Oy 151 GACTCGAGAGCATGATTTGGTATTCAGACACCAATTCAGACGGCATGTTGATTTCAA 92
Db 421 AlaTrpArgArgValLeuAlaGluValAspArgAspAsnAspGlyGlnValAspPheGln 440
Oy 91 GAATTCAAAACATGATGTTA 71
Db 441 GluPheGlnGlnMetLeuLeu 447

```

RESULT 13
08-668-416-2

GENERAL INFORMATION:

APPLICANT: Tomley, Fiona M.
APPLICANT: Dunn, Paul P. J.

APPLICANT: Bumstead, Janene M
APPLICANT: Vermeulen, Arno N

TITLE OF INVENTION: Coccidiosis poultry vaccine

```

;
; NUMBER OF SEQUENCES: 10
;
; CORRESPONDENCE ADDRESS:
;

```

ADDRESS: AKZO NO. 5843/22e1 Patent
STREET: 1300 Piccard Drive, Suite 206

CITY: Rockville
STATE: Maryland

COUNTRY: U
ZTD. 30950

FILE: 2000
COMPUTER READABLE FORM:

```

; MEDIUM TYPE: floppy disk
;
; COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS.
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: INS/08/668 416

DECLASSIFICATION
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION:

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/464,164
;; FILING DATE: June 2, 1995

ATTORNEY/AGENT INFORMATION
NAME: Gormley Mary

REGISTRATION NUMBER: 34,409

TELECOMMUNICATION INFORMATION;
TELEPHONE: (301) 258-5200

INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:

LENGTH: 456 amino acids
TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-668-416-2

Alignment Scores

Pred. No.:

score:	43.50	maximal:	51
Percent Similarity:	41.18%	Conservative:	26

Best Local Similarity:	27.2%	Mismatches:	39
Query Match:	13.71%	Indels:	51

DB:	2	Gaps:	6
-----	---	-------	---

US-10-021-323-13 (1-609) x US-08-668-416-2 (1-456)

US-10-021-323-13 (1-609) X US-08-668-416-2 (1-456)

565 ATCTCTTTGGTGTAAATGTCCGCCCTTAGTAGACCGAC--TTGCAACGGCTATTC 509

Db 290 LeuTyrMetGlySerIysLeuThrThrAsnGluGluThrAspGluLeuAsnLysIlePhe 309

508 GAGAGCTGACAGAGATGACAGATGACGCTTGGTCTGAGAGAGCTGAAATTCGTTGCTC 449

228

310 GANDYSMECASPUSABNGIYASPRGTYGNIWEASPCYSQINQIUEU..... 322

448 CAGAGATCGGGTCTGTCCAATTACGCTTGAAGATTGGAGCCCTAGTGGGAAAACCA 389

Db 326 -----MetGIuGIYrValGIuLeuMetLysLeuLysGIuGIuAsp 3335

388 TGTTTGAACCTTGATGAAATCTTGTCTTTTATGAATCCATCTGAACCACTGGACAT 329

```

Db      340 ValServalleuaplybSeraIaIe----- 348
      |||||:::

```

328 GCGCGTGA CGA AGA GGA GGA GGA GGA ATTTGGTCA TTTCA CGGC-----GGC 284

366 376 386 396 406 416 426 436 446 456 466 476 486 496 506 516 526 536 546 556 566 576 586 596 606 616 626 636 646 656 666 676 686 696 706 716 726 736 746 756 766 776 786 796 806 816 826 836 846 856 866 876 886 896 906 916 926 936 946 956 966 976 986 996

DB 349 -----GILMILL

283 GGTGAAGAAGACAGT

APPLICANT: Takezawa, D.
APPLICANT: Han, T. J.
TITLE OF INVENTION: Control of Growth and Development of
TITLE OF INVENTION: Potato Plants
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Harness, Dickey & Pierce
STREET: P.O. Box 828
CITY: Bloomfield Hills
STATE: MI
COUNTRY: USA
ZIP: 48303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/100,874
FILING DATE: July 30, 1993
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Dean F.
REFERENCE/DOCKET NUMBER: 7555-00004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (313) 641-1600
TELEFAX: (313) 641-0270
TELEX: 287637 Harness UR
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 149 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-100-874-2

Alignment Scores:
Pred. No.: 1,62e-08 Length: 149
Score: 145.00 Matches: 43
Percent Similarity: 49.67% Conservative: 33
Best Local Similarity: 28.10% Mismatches: 59
Query Match: 13.67% Indels: 18
DB: 1 Gaps: 4

US-10-021-323-13 (1-609) x US-08-100-874-2 (1-149)

526 GACTTGACGCGATTCGAGAACCTGACAGAAAGATGAGATGCTTGGTGGAG 467
12 Gluhelysgluvalapheserleupheaplsyaspgllycyvillerthrlye 31
466 GAGCTGAATGCTGCTCCAGAGATCGGCTGTCCAAATTCACCTTGAA---GAATTG 410
32 Gluhelythrvalmetarserleugly-----glnanprothrgluvalglu 49
409 GAGCCCTTAGTGGAAAAACATGTTGAATGATTAATCTGTTTAAATGATCC 350
50 Glnaspmetclleserlgluvalaaspalnanglythrilleasppheproglu 69
349 ATTCGAAACCACTGCGACATGCTGAGCAAGAGAGAGAGAAATTCGATTCAC 290
70 Leuasnleuvalaarglyvmetlyasphrthraspergluglu----- 85
289 GCGCGCGTGAAGAAGACAGTGCCTTGCAGAGCTTTAAAGTGTGATGAT 230
86 -----leulysgluvalaphelysvalphaspllyasp 96
229 GGGGATGGTGGGGGATGTGAGAGCTTGAATCGTGTGGAGAACTGGGATGTA 170
97 Glnasnglyphelieseralaalgluvalarghlsvalmetthrnsleugly----- 114
169 GGTGAAATAGTGAAGAAAGCTGACAGAGCATGATTTGATTCAGACCAATTCAGAC 110

Db 115 Gluhelythrvalasplgluvalaaspgluvalarggluvalaasppllyasp 134
Cy 109 GCGATGCTGATTTTCAAGATTCAGAAACATGATGTTA 71
Db 135 Glylnvalasntyrngluvalaphelysvalargmetmetleu 147

RESULT 15

US-08-963-409-5
Sequence 5, Application US/08963409
Patent No. 6046315

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

APPLICANT: Shah, Purvi

APPLICANT: Corley, Neil C.

TITLE OF INVENTION: DISEASE ASSOCIATED CALMODULIN PROTEIN

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Dr.

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette

COMPUTER: IBM compatible

OPERATING SYSTEM: DOS Windows Version 2.0

SOFTWARE: FASTSEQ for windows

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/963,409

FILING DATE: Filed Herewith

PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0418 US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 152 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:
LIBRARY: Genbank

CLONE: 825635

US-08-963-409-5

Alignment Scores:
Pred. No.: 1.87e-08 Length: 152
Score: 144.50 Matches: 41
Percent Similarity: 47.80% Conservative: 35
Best Local Similarity: 25.79% Mismatches: 54
Query Match: 13.62% Indels: 29
DB: 3 Gaps: 5

US-10-021-323-13 (1-609) x US-08-963-409-5 (1-152)

526 GACTTGACGCGATTCGAGAACCTGACAGAAAGATGAGATGCTTGGTGGAG 467
12 Gluhelysgluvalapheserleupheaplsyaspgllythrillerthrlye 31
466 GAGCTGAATGCTGCTCCAGAGATCGGCTGTCCAAATTCACCTTGAA---GAATTG 425
32 Gluhelythrvalmetarserleuglyglnasprothrgluvalglu 51
424 AGCTTGAAGAAATG---GAGCCCTTAGTGGAGAAACATGTTGAATGATGAA 371
52 Metileasngluvalaasppllyaspplleuprogllyasnanglythrilleasppheproglu 71

Mon Jun 30 09:38:05 2003

us-10-021-323-13.rtf

Page 10

```

OY 370 TTTCTGTTCTTTATGATTCACCTCGAAACCCCTGCACATGCTGGTGCAGAAAGAG 311
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 72 PheLeu-----ThiMeMetChaIArglySMeTlySaBpThrAsp 84
OY 310 GAGAGGAATTGTCATTCAACGGCGCGGTGAAGAAAGACAGTCACCTTGGCAAGGCT 251
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 85 SerGIuGIu-----GluIleArgGIuAla 92
OY 250 TTTAAGGTGTGACTTGATGATGGATGGATGGATGATGAGAGCTTGAAATACGTG 191
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 93 PheArgValAPheAspIysAspGlyIleSerLlaIaGIuLysuAArgHLeVal 112
OY 190 CTGGGAGAAGCTGGGTATGTGAGGTGAATAATAGTGAAAAGACTGCAGAGACATGATTGG 111
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 113 MetThrAsnLeuGly-----GluYsLeuThrAspGIuGIuValAspGIuMetIleArg 130
OY 130 TATTACACACCAATTCAGACGGCATGGTGTGATTTTCAAGAAATTCAAAAACATGATG 74
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 131 GluAlaAspIleAspGlyAspGlyGIuValAsnTyrGIuGIuIuIuHeValGIuMetMet 149

```

Search completed: June 24, 2003, 14:06:42
Job time : 19.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 24, 2003, 14:02:28 ; Search time 28.5 Seconds
(without alignments)
4624.417 Million cell updates/sec

Title: US-10-021-323-13

Perfect score: 1100
Sequence: 1 ggcgaatgcaatcaactttc.....tgaagctntacaatcaagg 609

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 41779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 835558

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ n2p.model -DEV=xip
-O/cgn2_1/USPTO.spool/US10021323/runat 24062003 102236 20598/app query.fasta_1.775
-DB=Published Applications AA -QFMT=fscan -SUFFIX=radb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonsum62
-TRANS=human40.csl -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10021323 @CGN 1.1 17 @runat 24062003 102236 20598
-NCPU=6 -ICPU=3 -NO MAP -LARGESURVEY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-Fgapop=6 -Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications AA:*

1: /cgn2_6/prodata/2/pubppaa/US08_NEM_PUB.pep:*
2: /cgn2_6/prodata/2/pubppaa/PCR_NEM_PUB.pep:*
3: /cgn2_6/prodata/2/pubppaa/US06_NEM_PUB.pep:*
4: /cgn2_6/prodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/prodata/2/pubppaa/US07_NEM_PUB.pep:*
6: /cgn2_6/prodata/2/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/prodata/2/pubppaa/PCrUS_PUBCOMB.pep:*
8: /cgn2_6/prodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/prodata/2/pubppaa/US09_NEM_PUB.pep:*
10: /cgn2_6/prodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/prodata/2/pubppaa/US10_NEM_PUB.pep:*
12: /cgn2_6/prodata/2/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/prodata/2/pubppaa/US60_NEM_PUB.pep:*
14: /cgn2_6/prodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	239.5	22.6	259	9	US-09-847-208-45
C 2	154	14.5	159	10	US-09-910-071-5
C 3	152.5	14.4	160	9	US-10-192-806-1
C 4	147	13.9	642	9	US-09-554-000-2

C 5	147	13.9	652	9	US-09-554-000-4	Sequence 4, Appl1
C 6	144	13.6	642	9	US-09-554-000-6	Sequence 8, Appl1
C 7	144	13.6	656	9	US-09-554-000-8	Sequence 8, Appl1
C 8	143	13.5	408	9	US-09-988-462-21	Sequence 21, Appl1
C 9	143	13.5	464	9	US-09-988-462-22	Sequence 22, Appl1
C 10	142	13.4	416	9	US-09-989-025A-8	Sequence 8, Appl1
C 11	141	13.3	148	9	US-09-989-025A-4	Sequence 4, Appl1
C 12	136.5	12.9	142	10	US-09-910-071-4	Sequence 4, Appl1
C 13	132	12.4	142	9	US-09-988-462-24	Sequence 24, Appl1
C 14	127.5	12.0	146	9	US-10-157-031-295	Sequence 295, App
C 15	118	11.1	639	10	US-09-854-731-17	Sequence 17, Appl1
C 16	114.5	10.8	463	9	US-09-988-462-25	Sequence 25, Appl1
C 17	114	10.7	549	10	US-09-828-313-39	Sequence 39, Appl1
C 18	112	10.6	172	12	US-10-109-885-4	Sequence 4, Appl1
C 19	107	10.1	140	10	US-09-925-301-1486	Sequence 1486, Ap
C 20	102	9.3	633	9	US-10-086-464-11	Sequence 11, Appl1
C 21	101	9.5	201	10	US-09-925-297-714	Sequence 714, App
C 22	99.5	9.4	170	12	US-10-109-885-3	Sequence 3, Appl1
C 23	99	9.3	475	10	US-09-777-921A-4	Sequence 4, Appl1
C 24	98	9.2	300	10	US-09-810-672A-6	Sequence 6, Appl1
C 25	98	9.2	410	10	US-09-777-921A-5	Sequence 5, Appl1
C 26	98	9.2	477	10	US-09-777-921A-2	Sequence 2, Appl1
C 27	97.5	9.2	137	9	US-10-102-806-658	Sequence 658, App
C 28	97.5	9.2	331	9	US-09-961-403-11	Sequence 11, Appl1
C 29	97.5	9.2	331	10	US-09-768-840-3	Sequence 3, Appl1
C 30	97.5	9.2	331	10	US-09-847-809A-5	Sequence 5, Appl1
C 31	96.5	9.1	335	10	US-09-768-840-4	Sequence 4, Appl1
C 32	95	8.6	957	9	US-10-155-400-1	Sequence 1, Appl1
C 33	94	8.9	342	10	US-09-777-921A-6	Sequence 6, Appl1
C 34	94	8.9	379	10	US-09-810-672A-4	Sequence 4, Appl1
C 35	94	8.9	352	10	US-09-810-672A-2	Sequence 2, Appl1
C 36	94	8.5	730	10	US-09-801-368-126	Sequence 126, App
C 37	94	8.5	1228	9	US-09-917-384-1	Sequence 1, Appl1
C 38	94	8.5	1228	9	US-09-917-383-1	Sequence 1, Appl1
C 39	92.5	8.7	189	9	US-09-764-891-3897	Sequence 3897, Ap
C 40	92.5	8.4	731	9	US-10-086-464-17	Sequence 17, Appl1
C 41	92	8.7	57	10	US-09-834-765-770	Sequence 770, App
C 42	91	8.6	133	9	US-10-157-031-285	Sequence 285, App
C 43	90	8.2	116	9	US-09-864-761-40290	Sequence 40290, A
C 44	90	8.2	178	9	US-10-157-031-171	Sequence 171, App
C 45	89.5	8.1	1506	9	US-10-142-650-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-09-847-208-45
Sequence 45, Application US/09847208
Publication No. US20030082190A1
GENERAL INFORMATION:
APPLICANT: Saxon, Andrew
APPLICANT: Zhang, Ke
TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
FILE REFERENCE: DC67, 002A
CURRENT FILING DATE: 2001-05-01
NUMBER OF SEQ ID NOS: 177
SOFTWARE: PASCSEQ for Windows Version 4.0
SEQ ID NO 45
LENGTH: 205
TYPE: PRT
ORGANISM: Betula verrucosa (White birch). (Betula pendula)
US-09-847-208-45

Alignment Scores:
Pred. No.: 2.62e-16
Score: 239.50
Percent Similarity: 57.93%
Best Local Similarity: 38.41%
Query Match: 22.57%
DB: 9
Gaps: 6

US-10-021-323-13 (1-609) x US-09-847-208-45 (1-205)

```

Oy      538  CTTAGGAAGACCGACATGGACGCGATTCGGAAGAGCTGCACGAATATGAGATGGCTTC 479
Db      36  LeuAenThrlEuArglEuArgrglrIlePheApleuPheAplYpAsnSerApsglYIle 55
Oy      478  GTTAGTCTGGAGAGAGCTGAATTGGTTGCTCCAGAAATCGGGTCTTCCAATTACGCTT 419
Db      56  IleThrValAspIuAserThrArgAlaLeuAenLeuEngly---LeuGluThrAspLeu 74
Oy      418  GAAGAATTGGAGCCCTTAGTG-----GGAAAACATGTTTGAACCTG 377
Db      75  SerGluLeuGluSerThrValIysSerPheThrArgGluIysAenIleGlyLeuGlnPhe 94
Oy      376  GARGAATTCCTGTCTTTATGAATCCATCGAGACCCA---CTGGACATGGGNGGT--- 323
Db      95  GluAspPheIleSerLeuHieGlnSerLeuAenApsSerYrPheAlaArgYlGlyGlu 114
Oy      322  GACGAAGAGAGAGAGAGAAATGTGCATTCACGCGCGGTGAAGAAGAACAGACGTAC 263
Db      115  AspGluApsApsAenIuGluApsMeArGlyrSserIleLeuSerGlnIuAlaAsp 133
Oy      262  CTTGCGAAGGCTTTTAAAGTGTGTCATTGAATGGAGATGGGTGGGGGAGTGTAGAG 203
Db      135  SerPheGlyYlYheIysValPheApsGluApsGlyApsGlyTYrIleSerAlaArgGlu 154
Oy      202  CTTGATATCGTCTGTGGAGAGACTGGGTATGTGAGGTGAATAATGTGAAAAAGAC---TGC 146
Db      155  LeuGlnMetValIeuGlyIysLeuEnglyPhe---SerGluGlySerGluIleApsArgVal 173
Oy      145  AGAGAGCATGTTTGATTAACGACACCAATTCAGACCGGATCGGTGAATTTCAAGAATTC 86
Db      174  GluIysMetIleValSerValApsSerApsApsApsGlyArgValApsPhePheGluPhe 195
Oy      85  AAAAACAATGATG 74
Db      194  LysApsMetMet 197

```

```

RESULT 2
US-09-910-071-5
Sequence 5, Application US/09910071
Patent No. US20020116146A1
GENERAL INFORMATION:
APPLICANT: Tomikawa, Mayumi
APPLICANT: Aikawa, Seichi
APPLICANT: Matsuzawa, Fumiko
TITLE OF INVENTION: Method and Apparatus for Extracting and Evaluating Mutually Similar
TITLE OF INVENTION: Portions in One-Dimensional Sequences in Molecules and/or Three-
DIME
TITLE OF INVENTION: Structures of Molecules
FILE REFERENCE: 522.1921D2
CURRENT APPLICATION NUMBER: US/09/910,071
CURRENT FILING DATE: 2001-07-23
PRIOR APPLICATION NUMBER: US 08/014,867
PRIOR FILING DATE: 1993-02-08
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 159
TYPE: PRT
ORGANISM: unknown
FEATURE:
OTHER INFORMATION: Figure 23B
US-09-910-071-5

```

Alignment Scores:	
Pred. No.:	1,79e-07
Score:	154.00
Percent Similarity:	45.22%
Best Local Similarity:	28.03%
Query Matchn:	14,51*
OB:	10
Length:	156
Matches:	44
Conservative:	27
Mismatches:	26
Indels:	2
Gaps:	5

US-10-021-323-13 (1-609) X US-09-910-071-5 (1-159)

```

Oy 526 GAGCTTCAACGGCGTATTCGAAAGCTCCAGCAAGAAATGAGATGGCTTCGTTAGCTGAG 467
Db   :::      |||      |||      |||      |||      |||      |||      |||
Oy 19  GluPheLysAlaAlaPheAspMetPheAspAlaAspGluGluAspAlaIleSerThrLys 38
Db   ::::      ::::      ::::      ::::      ::::      ::::      ::::      ::::
Oy 466 GAGCTGAATTTGGTGTCTCCAGAGAAATCGGGTGTCTCCATTACAGCTTTGAAGAAATTTGAG 407
Db   ::::      ::::      ::::      ::::      ::::      ::::      ::::      ::::
Oy 39  GluLeuGluYthrValMetArgMetLeuGluGln--AsnProThrLysGluGluLeuAsp 57
Db   ::::      ::::      ::::      ::::      ::::      ::::      ::::      ::::
Oy 406 CCTTAGTG-----GGAACAACATGTTGAATCTTGATGAATTTCTTG 365
Db   ::::      ::::      ::::      ::::      ::::      ::::      ::::      ::::
Oy 58  AlaIleIleGluGluValAspGluAspGlySerGlyThrIleAspPheGluPheLys 77
Db   ::::      ::::      ::::      ::::      ::::      ::::      ::::      ::::
Oy 364 TTCCTTTATGAATTCATCTCGAACCCAGCTGGCAGCATGGTGTGACGAAGAGAGAGAGAG 305
Db   ::::      ::::      ::::      ::::      ::::      ::::      ::::      ::::
Oy 78  ValMetValArgGluMetLysGluAsp--AlaLysGlyLysSerGluGluGlu----- 94
Db   ::::      ::::      ::::      ::::      ::::      ::::      ::::      ::::
Oy 304 GAATTGGTCATTACGGCGCGCGGTGAAGAAGAACAATGACCTTGGGAAGGCTTTAAA 245
Db   -----LeuAlaAspCysPheArg:::
Oy 95  -----LeuAlaAspCysPheArg:::
Db   -----LeuAlaAspCysPheArg:::
Oy 244 GTGTTGACTTGAATGGAGATGGATGGTGGGGAGATGAGAGCTTGAATACGTGCTGGA 185
Db   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 101 IlePheAspLysAsnAlaAspLysPheIleAspIleGluGluLeuGluGluIleLeuArg 120
Db   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 184 AGACTGGGTATGTGAGCTGAATAATGTGGAAGAAGCTCCAGAGACATGATTGGTATTAC 125
Db   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 121 AlaThrGly-----GluHisValThrGluGluAspIleGluAspLeuMetLysAspSer 158
Db   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 124 GAGACCAATTCAGAGCGCATGGTTGATTTTCAAGATTCCAAAAACATGATG 74
Db   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 139 AspLysAsnAsnAspGlyArgLysAspPheAspGluPheLeuLysMetMet 155
Db   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

, RESULT 3
, US-10-192-806-1
, Sequence 1, Application US/10192806
, Publication No. US20030083255A1
, GENERAL INFORMATION:
, APPLICANT: Thorn, R.
, APPLICANT: Lanzer, M.
, APPLICANT: Moses, M.
, APPLICANT: Wiederschain, D.
, TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS COMPRISING
, TITLE OF INVENTION: TROPONIN SUBUNITs, FRAGMENTS AND HOMOLOGS THEREOF AND
, TITLE OF INVENTION: METHODS OF THEIR USE TO INHIBIT ANGIOGENESIS
, FILE REFERENCE: 9657-042-999
, CURRENT APPLICATION NUMBER: US/10/0192,806
, CURRENT FILING DATE: 2002-07-09
, PRIOR APPLICATION NUMBER: 09/612,421
, PRIOR FILING DATE: 2000-07-07
, PRIOR APPLICATION NUMBER: 09/268,274
, PRIOR FILING DATE: 1999-03-15
, PRIOR APPLICATION NUMBER: 08/961,264
, PRIOR FILING DATE: 1997-10-30
, PRIOR APPLICATION NUMBER: 08/602,941
, PRIOR FILING DATE: 1996-02-16
, NUMBER OF SEQ ID NOS: 20
, SOFTWARE: FastSeq for Windows Version 3.0
, SEQ ID NO 1
, LENGTH: 160
, TYPE: PRT
, ORGANISM: Homo sapiens
, US-10-192-806-1

```

```

Pairing Score:      2.57e-07      Length: 166
Pred. No.:          152.50        Matches: 29
Score:              45.86         Conservative: 23
Percent Similarity: 27.39%        Mismatches: 60
Best Local Similarity: 14.37%     Indels: 25
Query Match:        9             Gaps: 4
DB:

```


US-10-021-323-13 (1-609) x US-10-192-806-1 (1-160)

QY 526 GACTTGCAACGCGATTTCGAGAACTCGACAGATGAGATGCTTCGTTAGCTTGAG 467
 Db 19 GluphelysalaaIaPheAspMetPheAspAlaAspIlyglYAspIleSeVallys 38
 QY 466 GAGCTGAATGGTTGCTCCAGAGAACTGCTCCAAATTCAGCTTGAAGATTTGAG 407
 Db 39 GlulduelglYthrValMetArgMetLeuglYglInthrPro---ThrlySgluGlulduasp 57
 QY 406 CCCTAGATG-----GGAAGAACATGTTTGAACCTTGAGATTCCTTG 365
 Db 58 AlalIleelgluIgluValaAspIlyserIlyThrIleAspPhegluIgluPheleu 77
 QY 364 TTCTTTATGAATTCATCTCGAACCCACTGACATGCTGTGTGACGAAGAGAGAGAG 305
 Db 78 ValMetMetValArgIglMetIlySgluAspAlaIlySglYserIgluIglu----- 95
 QY 304 GAATGTGTCATTCACGGCGGGCGGAGAGAGAACAGATGACCTTGGAAGGCTTTAA 245
 Db 96 -----LeuAlaIgluCYsPhearg 101
 QY 244 GTGTTGACTGAATGAGATGGATGGGAGATGAGAGACTTGAATACGCTGGA 185
 Db 102 IllePheAspArgAsnAlaAspGlyTyrIleAspProgluIgluAlaIgluIlePhearg 121
 QY 184 AGATGGATGTAGAGGTGAAATATGTAAGAAAGACTGCAGAGACATGATTTGATTTAC 125
 Db 122 AlaserIly-----GlulIseValThrAspIgluIleIgluSerLeuMetIlyAspIly 139
 QY 124 GACACCAATTCAGACGCGATGATTTTCAGAAATTCAAAAATCATGATG 74
 Db 140 AspIlyAsnAsnAspGlyArgIleAspPheAspIgluPheIuIyMetMet 156

RESULT 4
 US-09-554-000-2
 ; Sequence 2, Application US/09554000
 ; Patent No. US20020165364A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tsien, Roger Y.
 ; APPLICANT: Miyawaki, Akiyoshi
 ; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
 ; TITLE OF INVENTION: DETECTION OF ANALYTES
 ; FILE REFERENCE: 07257/042001
 ; CURRENT APPLICATION NUMBER: US/09/554,000
 ; CURRENT FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: 08/818,252
 ; PRIOR FILING DATE: 1997-03-14
 ; NUMBER OF SEQ ID NOS: 56
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 642
 ; TYPE: PRT
 ; ORGANISM: Aequorea victoria
 ; US-09-554-000-2

Alignment Scores:
 Pred. No.: 1,46e-06 Length: 642
 Score: 147.00 Matches: 41
 Percent Similarity: 50.66% Conservative: 36
 Best Local Similarity: 26.97% Mismatches: 57
 Query Match: 13.85% Indels: 18
 DB: 9 Gaps: 4

US-10-021-323-13 (1-609) x US-09-554-000-2 (1-642)

QY 526 GACTTGCAACGCGATTTCGAGAACTCGACAGATGAGATGCTTCGTTAGCTTGAG 467
 Db 241 GluphelysalaaIaPheAspMetPheAspAlaAspIlyglYAspIleSeVallys 260
 QY 466 GAGCTGAATGGTTGCTCCAGAGAACTGCTCCAAATTCAGCTTGAAGATTTGAG 410
 Db 261 GlulduelglYthrValMetArgMetLeuglYglInthrProThrIgluIglu 278

QY 409 GAGCCCTAGTGGGAAAACCATGTTGAACCTTGATGAATTCCTGTTTATGATCC 350
 Db 279 GluAspMetIleAsnIgluValaAspIlyAsnIglYThrIleTyrPheProgluPhe 298
 QY 349 ATCTCGAACCCACTGCGACATGCTGTGTCAGAGAGAGAGAGAGATTCATTCAC 290
 Db 299 LeuThrMetMetAlaArgIlyMetIlyAspThrAspSerIgluIglu----- 314
 QY 289 GCGCGCGGTGAAGAGAGACACTGACCTTCCAGAGGCTTTAAAGTGTTCGATTCAT 230
 Db 315 -----IleArgIgluAlaPheArgValPheAspIlyAsp 325
 QY 229 GGGATGCGTTGGGGGATGATGAGACTTGAATACCTGCTGGGAGACCTGGGATTCGTA 170
 Db 326 GlyAsnIlyTyrIleSerAlaIgluIgluValaAspIlyValMetThrAsnIgluY----- 343
 QY 169 GGTGAAAATATGTAAGAAAGACTGCAGAGACATGATTTGATTCACACCAATTCAGAC 110
 Db 344 GlulYsleuThrAspIgluIgluValaAspIgluMetIleArgIgluAlaAspIleAspIlyAsp 363
 QY 109 GCGATGTTGATTTTCAGAAATTCAAAAATCATGATG 74
 Db 364 GlYlgluValAsnTyrIgluIgluPheValIgluMetMet 375

RESULT 5
 US-09-554-000-4
 ; Sequence 4, Application US/09554000
 ; Patent No. US20020165364A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tsien, Roger Y.
 ; APPLICANT: Miyawaki, Akiyoshi
 ; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
 ; TITLE OF INVENTION: DETECTION OF ANALYTES
 ; FILE REFERENCE: 07257/042001
 ; CURRENT APPLICATION NUMBER: US/09/554,000
 ; CURRENT FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: 08/818,252
 ; PRIOR FILING DATE: 1997-03-14
 ; NUMBER OF SEQ ID NOS: 56
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 652
 ; TYPE: PRT
 ; ORGANISM: Aequorea victoria
 ; US-09-554-000-4

Alignment Scores:
 Pred. No.: 1,47e-06 Length: 652
 Score: 147.00 Matches: 41
 Percent Similarity: 50.66% Conservative: 36
 Best Local Similarity: 26.97% Mismatches: 57
 Query Match: 13.85% Indels: 18
 DB: 9 Gaps: 4

US-10-021-323-13 (1-609) x US-09-554-000-4 (1-652)

QY 526 GACTTGCAACGCGATTTCGAGAACTCGACAGATGAGATGCTTCGTTAGCTTGAG 467
 Db 241 GluphelysalaaIaPheAspMetPheAspAlaAspIlyglYAspIleSeVallys 260
 QY 466 GAGCTGAATGGTTGCTCCAGAGAACTGCTCCAAATTCAGCTTGAAGATTTGAG 410
 Db 261 GlulduelglYthrValMetArgMetLeuglYglInthrProThrIgluIglu 278
 QY 409 GAGCCCTAGTGGGAAAACCATGTTGAACCTTGATGAATTCCTGTTTATGATCC 350
 Db 279 GluAspMetIleAsnIgluValaAspIlyAsnIglYThrIleTyrPheProgluPhe 298
 QY 349 ATCTCGAACCCACTGCGACATGCTGTGTCAGAGAGAGAGAGAGATTCATTCAC 290
 Db 299 LeuThrMetMetAlaArgIlyMetIlyAspThrAspSerIgluIglu----- 314
 QY 289 GCGCGCGGTGAAGAGAGACACTGACCTTCCAGAGGCTTTAAAGTGTTCGATTCAT 230
 Db 315 -----IleArgIgluAlaPheArgValPheAspIlyAsp 325
 QY 229 GGGATGCGTTGGGGGATGATGAGACTTGAATACCTGCTGGGAGACCTGGGATTCGTA 170
 Db 326 GlyAsnIlyTyrIleSerAlaIgluIgluValaAspIlyValMetThrAsnIgluY----- 343
 QY 169 GGTGAAAATATGTAAGAAAGACTGCAGAGACATGATTTGATTCACACCAATTCAGAC 110
 Db 344 GlulYsleuThrAspIgluIgluValaAspIgluMetIleArgIgluAlaAspIleAspIlyAsp 363

QY 289 GCGCGCGTGAAGAAGACAGTGCCTTGGCAAGCTTTAAAGTCTTGACTTGAT 230
Db 315 -----1learglualapheargvalpheasplyasp 325
QY 229 GGGATGGGTGGGGGATGTGAGAGCTTGATACGTGCTGGGAAGACTGGATGTGA 170
Db 326 Glysenglytryllesealalaglueuarghsvalmetthrashleugly----- 343
QY 169 GGTGAATAAGTGAAGAAGCTGAGAGCATGATTTGATTAACACCAATTGACAC 110
Db 344 Glulysleuthrarpglugluvalaspdlumetlleargglualaspdlasp 363
QY 109 GGCATGGTGAATTTCAAGATTCAAAAACATGATG 74
Db 364 Glylnvalaenlyrglugluhphevalglmetmet 375

RESULT 6
US-09-554-000-6
Sequence 6, Application US/09554000
Patent No. US20020165364A1
GENERAL INFORMATION:
APPLICANT: Miyawaki, Atsushi
TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
FILE REFERENCE: 07257/042001
CURRENT APPLICATION NUMBER: US/09/554,000
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: 08/818,252
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 642
TYPE: PRF
ORGANISM: Aequorea victoria
US-09-554-000-6

Alignment Scores:
Pred. No.: 2,996-06 Length: 642
Score: 144.00 Matches: 40
Percent Similarity: 50.66% Conservative: 37
Best Local Similarity: 26.32% Mismatches: 57
Query Match: 13.57% Indels: 18
Gaps: 4
DB: 9

US-10-021-323-13 (1-609) x US-09-554-000-6 (1-642)

526 GACTTGCAACCGGATTCGAGAGCTCGACAAAGATGAGATGCTTGTGATGAG 467
Db 241 Glupheylsglualapheserleupheasplyaspdllythrillethrlyrs 260
QY 466 GAGCTGAATGGTCTCCAGAGATCGGTCTGCCAATTCAGCTTGA--GAATG 410
Db 261 Glueuglythrvalmetargserleugly-----Glnaenprothrglualaglueu 278
QY 409 GAGCCCTTAGTGGGAAAACCATGTTGAATCTGATGAATCTTCTTTATGATCC 350
Db 279 Glnaspmetlleaenclvalaspdlaspdlaspdlaspdlaspdlaspdlasp 298
QY 349 ATCTGGAACCCACTGACAGATGCTGTGACGAGAGAGAGAGAGATGTCATTCAC 290
Db 289 Leuthrmetlealarglysmetlyasprhrasprserglugluu----- 314
QY 289 GCGCGCGTGAAGAAGACAGTGCCTTGGCAAGCTTTAAAGTCTTGACTTGAT 230
Db 315 -----1leargglualapheargvalpheasplyasp 325
QY 229 GGGATGGGTGGGGGATGTGAGAGCTTGATACGTGCTGGGAAGACTGGATGTGA 170
Db 326 Glysenglytryllesealalaglueuarghsvalmetthrashleugly----- 343
QY 169 GGTGAATAAGTGAAGAAGCTGAGAGCATGATTTGATTAACACCAATTGACAC 110

Db 344 Glulysleuthrarpglugluvalaspdlumetlleargglualaspdlasp 363
QY 109 GGCATGGTGAATTTCAAGATTCAAAAACATGATG 74
Db 364 Glylnvalaenlyrglugluhphevalglmetmet 375

RESULT 7
US-09-554-000-8
Sequence 8, Application US/09554000
Patent No. US20020165364A1
GENERAL INFORMATION:
APPLICANT: Tsien, Roger Y.
TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
FILE REFERENCE: 07257/042001
CURRENT APPLICATION NUMBER: US/09/554,000
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: 08/818,252
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 656
TYPE: PRF
ORGANISM: Aequorea victoria
US-09-554-000-8

Alignment Scores:
Pred. No.: 3,016-06 Length: 656
Score: 144.00 Matches: 40
Percent Similarity: 50.66% Conservative: 37
Best Local Similarity: 26.32% Mismatches: 57
Query Match: 13.57% Indels: 18
Gaps: 4
DB: 9

US-10-021-323-13 (1-609) x US-09-554-000-8 (1-656)

QY 526 GACTTGCAACCGGATTCGAGAGCTCGACAAAGATGAGATGCTTGTGATGAG 467
Db 257 Glupheylsglualapheserleupheasplyaspdllythrillethrlyrs 276
QY 466 GAGCTGAATGGTCTCCAGAGATCGGTCTGCCAATTCAGCTTGA--GAATG 410
Db 277 Glueuglythrvalmetargserleugly-----Glnaenprothrglualaglueu 294
QY 409 GAGCCCTTAGTGGGAAAACCATGTTGAATCTGATGAATCTTCTTTATGATCC 350
Db 295 Glnaspmetlleaenclvalaspdlaspdlaspdlaspdlaspdlaspdlasp 314
QY 349 ATCTGGAACCCACTGACAGATGCTGTGACGAGAGAGAGAGATGTCATTCAC 290
Db 315 Leuthrmetlealarglysmetlyasprhrasprserglugluu----- 330
QY 289 GCGCGCGTGAAGAAGACAGTGCCTTGGCAAGCTTTAAAGTCTTGACTTGAT 230
Db 331 -----1leargglualapheargvalpheasplyasp 341
QY 229 GGGATGGGTGGGGGATGTGAGAGCTTGATACGTGCTGGGAAGACTGGATGTGA 170
Db 342 Glysenglytryllesealalaglueuarghsvalmetthrashleugly----- 359
QY 169 GGTGAATAAGTGAAGAAGCTGAGAGCATGATTTGATTAACACCAATTGACAC 110
Db 360 Glulysleuthrarpglugluvalaspdlumetlleargglualaspdlasp 379
QY 109 GGCATGGTGAATTTCAAGATTCAAAAACATGATG 74
Db 380 Glylnvalaenlyrglugluhphevalglmetmet 391

RESULT 8
US-09-986-462-21

```

Db 264 IleAspIysAspAsnSerGlyThrIleThrLeuMetGluLeuLysHISGlyLeuAlaLys 283
QY 442 ATCGGGTCTGTCCAAATTCAGCCTTGAAGAATTTGAGCCCTTAGT----- 398
Db 284 HIGLYPro---LysLeuSerAspSerGluMetGluLysLeuMetGluAlaAlaAspAla 302
QY 397 ---GSAANAACATGTTTGAACCTTGATGAATTCCTGTTCTTTATGATCATCATCGAAC 344
Db 303 AspGlyAsnGlyLeuLeuLeuAspTyrAspGluPheVal----- 314
-QY 340 CCACTGCACATGTTGTGTCAGAAAGAGAGAGAGAGCAATTCATTCACGGCGCGGT 281
Db 315 -----ThraIaThrValHISMetAsnLys 322
QY 280 GAAGAAGAAAGACAGTGCACCTTCCGAAGCCTTTTAAAGTGTTTGACTGGAATGGGATGG 221
Db 323 LeuAspArgGluGluHISLeuTyrThrIaPheGlnTyrPheAspLysAspAsnSerGly 342
QY 220 TTGGGGGAGATGAGAGACTTGAATACGTCCTCGGAACACGCGGTATGTAGCGAAT 161
Db 343 TyrIleThrLysGluGluLeuGlnHISaLeuLysGluGlnGlyLeuTyrAspAlaAsp 362
QY 160 AGTGAAGAAAGCTGCAGAGCATGATTGTATTACGACACCAATTCAGACGGCATGTT 101
Db 363 LysIleLysAsp-----IleIleSerAspAlaAspSerAspAsnAspGlyArgIle 379
QY 100 GATTTCAGAATTCAAAAACATGATG 74
Db 380 AspTyrSerGluPheValaIaMetMet 388

RESULT 9
US-09-988-462-22
: Sequence 22, Application US/09988462
: Publication No. US20030046726A1
: GENERAL INFORMATION:
: APPLICANT: Kozziel, Michael G.
: Deasai, Nalin M.
: Lewis, Kelly S.
: Kramer, Vance C.
: Warren, Gregory W.
: Evola, Stephen V.
: Crossland, Lyle D.
: Wright, Martha S.
: Merlin, Ellis J.
: Launis, Karen L.
: TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
: INSECTICIDAL ACTIVITY IN MAIZE
: NUMBER OF SEQUENCES: 94
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Syngenta Biotechnology, Inc.
: STREET: 3054 Cornwallis Road
: City: Research Triangle Park
: STATE: NC
: COUNTRY: USA
: ZIP: 27709
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/988,462
: FILING DATE: 20-NO. US20030046726A1-2001
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 09/547,422
: FILING DATE: 11-APR-2000
: APPLICATION NUMBER: US 08/459,504
: FILING DATE: 02-JUN-1995
: APPLICATION NUMBER: US 07/951,715
: FILING DATE: 25-SEP-1992
: APPLICATION NUMBER: US 07/772,027
: FILING DATE: 04-OCT-1991

```


Publication No. US2003006874A1
GENERAL INFORMATION:
APPLICANT: MAKI, JUNICHI
TITLE OF INVENTION: METHOD FOR PRODUCING A BIOSENSOR PROTEIN CAPABLE OF REGULATING A
TITLE OF INVENTION: FLUORESCENCE PROPERTY OF GREEN FLUORESCENT PROTEIN, AND THE BIOS
TITLE OF INVENTION: PRODUCED BY THE METHOD
FILE REFERENCE: 216339US0
CURRENT APPLICATION NUMBER: US/09/989,025A
CURRENT FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: JP/2000-356047
PRIOR FILING DATE: 2000-11-22
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 148
TYPE: PRT
ORGANISM: Rattus norvegicus
US-09-989-025A-4
Alignment Scores:
d. No.: 3,91e-06 Length: 148
Score: 141.00 Matches: 42
Percent Similarity: 48.10% Conservative: 34
Best Local Similarity: 26.58% Mismatches: 52
Query Match: 13.29% Indels: 30
Gaps: 6
US-10-021-323-13 (1-609) x US-09-989-025A-4 (1-148)
QY 526 GACTTGCAACGGCGATTTCGAGAAGCTCGACAAAGATGAGATGCTTCTAGTCTGAG 467
DB 11 GIUPELYSGIUALAPheserleupheasplysaasplyThrllethrlyrs 30
QY 466 GAGCTGAATTGGTCTCTCCAGAGAAATCGGCTGTCTCAATTCAGCTTGA--GAATTG 410
DB 31 GIULSUGLYThrValMetArgserleugly-----GlnasnProthrGlnlaaglulau 48
QY 409 GAGCCCTTAGTG-----GGAAAACCATGTTGAATTCGATGATTC 368
DB 49 GlnaapMetlleasnGluValaaspIlaaspIlyasnGlyThrlleasphProgluphe 68
QY 367 TTGTTCTTTATGAATCCATCTCGAACCCATGTCGATGTCGATGACGAAGAGAGAG 308
DB 69 Leu-----ThrMetMetIlaArglySmetIlysaasphThraser 81
QY 307 GAGAATTGTCATTCACGGCGCGGTGAAGAAAGACAGTGCACCTTCGAAAGCTTT 248
DB 82 GIUGLU-----GluIleArgGluIlaPhe 89
QY 247 AAAGTGTTCGATTGATGATGGATGGGATGTCGAGAGCTTGAATACGTGCTG 188
DB 90 ArgValaPheasplysaasplyasnGlyTyrIleaserIlaaglulauArgHisValMet 109
QY 187 GGAAGACTGGGTATGTGAGAGCTGAATAATGAGTGCAGAGCATGATTTGGTAT 128
DB 110 Thrasnleugly-----GluIlyseuThrAspGluGluValaaspIleuIleargIlu 127
QY 127 TACGACACCAATTCAGAGCGCATGTTGATTTTCAAGATTCAAAACATGATG 74
DB 128 AlaaspIleasplysaasplyasnGlyGlnValasnTyrGluGluPheValaGlnMet 145
RESULT 12
US-09-910-071-4
Sequence 4, Application US/09910071
GENERAL INFORMATION:
APPLICANT: Tomikawa, Mayumi
APPLICANT: Aikawa, Seichi
APPLICANT: Matsuzawa, Fumiko
TITLE OF INVENTION: Method and Apparatus for Extracting and Evaluating Mutually Simi
TITLE OF INVENTION: Portions in One-dimensional Sequences in Molecules and/or Three-
FILE REFERENCE: 522.1921D2

CURRENT APPLICATION NUMBER: US/09/910,071
CURRENT FILING DATE: 2001-07-23
PRIOR APPLICATION NUMBER: US 08/014,867
PRIOR FILING DATE: 1993-02-08
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 142
TYPE: PRT
ORGANISM: unknown
OTHER INFORMATION: Figure 23A
US-09-910-071-4
Alignment Scores:
Pred. No.: 1,13e-05 Length: 142
Score: 136.50 Matches: 42
Percent Similarity: 47.47% Conservative: 33
Best Local Similarity: 26.58% Mismatches: 52
Query Match: 12.87% Indels: 31
Gaps: 6
US-10-021-323-13 (1-609) x US-09-910-071-4 (1-142)
QY 526 GACTTGCAACGGCGATTTCGAGAAGCTCGACAAAGATGAGATGCTTCTAGTCTGAG 467
DB 7 GIUPELYSGIUALAPheserleupheasplysaasplyThrllethrlyrs 26
QY 466 GAGCTGAATTGGTCTCTCCAGAGAAATCGGCTGTCTCAATTCAGCTTGA--GAATTG 410
DB 27 GIULSUGLYThrValMetArgserleugly-----GlnasnProthrGlnlaaglulau 44
QY 409 GAGCCCTTAGTG-----GGAAAACCATGTTGAATTCGATGATTC 368
DB 45 GlnaapMetlleasnGluValaaspIlaaspIlyasnGlyThrlleasphProgluphe 64
QY 367 TTGTTCTTTATGAATCCATCTCGAACCCATGTCGATGTCGATGACGAAGAGAGAG 308
DB 65 Leu-----ThrMetMetIlaArglySmetIlysaasphThraser 77
QY 307 GAGAATTGTCATTCACGGCGCGGTGAAGAAAGACAGTGCACCTTCGAAAGCTTT 248
DB 78 GIUGLU-----IleArgGluIlaPhe 84
QY 247 AAAGTGTTCGATTGATGATGGATGGGATGTCGAGAGCTTGAATACGTGCTG 188
DB 85 ArgValaPheasplysaasplyasnGlyTyrIleaserIlaaglulauArgHisValMet 104
QY 187 GGAAGACTGGGTATGTGAGAGCTGAATAATGAGTGCAGAGCATGATTTGGTAT 128
DB 105 Thrasnleugly-----GluIlyseuThrAspGluGluValaaspIleuIleargIlu 122
QY 127 TACGACACCAATTCAGAGCGCATGTTGATTTTCAAGATTCAAAACATGATG 74
DB 123 AlaasnIleasplysaasplyasnGlyGlnValasnTyrGluGluPheValaGlnMet 140
RESULT 13
US-09-988-462-24
Sequence 24, Application US/09988462
Publication No. US20030046726A1
GENERAL INFORMATION:
APPLICANT: Kozziel, Michael G.
Desai, Nalini M.
Lewis, Kelly S.
Kramer, Vance C.
Warten, Gregory W.
Evola, Stephen V.
Crosland, Lyle D.
Wright, Martha S.
Merlin, Ellis J.
Lauais, Karen L.
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
INSECTICIDAL ACTIVITY IN MAIZE

NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: Syngenta Biotechnology, Inc.
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/988,462
FILING DATE: 20-No. US20030046726A1-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/547,422
FILING DATE: 11-APR-2000
APPLICATION NUMBER: US 08/459,504
FILING DATE: 02-JUN-1995
APPLICATION NUMBER: US 07/951,715
FILING DATE: 25-SEP-1992
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: S-188051
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8587
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 142 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: Protein
LOCATION: 1..142
OTHER INFORMATION: /note="human calmodulin protein
sequence as shown in Figure 33."
SEQUENCE DESCRIPTION: SEQ ID NO: 24:
nr-09-988-462-24

-gment Scores:

Pred. No.: 3,31e-05 Length: 142
Score: 132.00 Matches: 41
Percent Similarity: 47.47% Conservative: 34
Best Local Similarity: 25.95% Mismatches: 53
Query Match: 12.44% Indels: 30
DB: 9 Gaps: 6

US-10-021-323-13 (1-609) x US-09-988-462-24 (1-142)

QY 526 GACTTGCAACGGCTATTGAGAGCTCGACAGAAATGAGATGGCTTCGTTAGCTGAG 467
DB 8 GlnPheLysGlnAlaPheSerLeuPheAspGlyAspGlyThrIleThrLys 27
QY 466 GAGCTGAATTCGTGCTCCAGAGATCGGCTGTCCTCAATTCAGCCTGAA---GAATG 410
DB 28 GlnLeuGlyThrValMetArgSerLeuGly-----GlnAsnProThrGlnAlaGlnLeu 45
QY 409 GAGCCCTTAGTG-----GGAAAACCATGTTGAATTCGATGGAATTC 368
DB 46 GlnAspMetIleAsnGlnValaAspGlyAsnGlyThrIleAspPheProGlnPhe 65
QY 367 TTGTTCTTTTATGATTCATCTCGAAACCACTGGCACATGCTGCTGATGCAAGAGAG 308
DB 367 TTGTTCTTTTATGATTCATCTCGAAACCACTGGCACATGCTGCTGATGCAAGAGAG 308

DB 66 Leu-----ThrMetAlaArgLysMetLysAspThrAspSer 78
QY 307 GAGAAATTCATTCACGGCGCGGTGAAGAAGACAGTACCTTGCGAAGCTTTT 248
DB 79 GlnGlu-----GlnIleArgGlnAlaPhe 86
QY 247 AAAGTGTTCATTCGAATGGGATGGCTGGGGGATGTGAGACCTTGAATACCTGCG 188
DB 87 ArgValLysAspLysAspGlyAsnGlyThrIleSerAlaAlaGlnLeuArgHisValMet 106
QY 187 GGAACACCTGGTATGAGAGTGAATAATGTGGAAGAAAGCTGAGAGCATGATTTGAT 128
DB 107 ThrAsnLeuGly-----GlnLysLeuThrAspGlnGlnValaAspGlnMetIleArgGln 124
QY 127 TACGACACCAATTCAGACGGCATGCTGATTTTCAAGATTCAAAACATGATG 74
DB 125 AlaAspIleAspGlyAspGlyGlnValaSerGlnGlnPheValGlnMet 142

RESULT 14

US-10-157-031-295

Sequence 295, Application US/10157031

Publication No. US20030108890A1

GENERAL INFORMATION:

APPLICANT: Baranova, A. V.

APPLICANT: Yankovsky, N. K.

APPLICANT: Kozlov, A. P.

APPLICANT: Lobashev, A. V.

APPLICANT: Krukovskaya, L. L.

TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequence

FILE REFERENCE: 2760-103

CURRENT APPLICATION NUMBER: US/10/157,031

CURRENT FILING DATE: 2002-05-30

NUMBER OF SEQ ID NOS: 415

SOFTWARE: Patent in version 3.1

SEQ ID NO 295

LENGTH: 146

TYPE: PR

ORGANISM: Homo sapiens

US-10-157-031-295

Alignment Scores:

Pred. No.: 9.79e-05 Length: 146
Score: 127.50 Matches: 41
Percent Similarity: 45.34% Conservative: 32
Best Local Similarity: 25.47% Mismatches: 67
Query Match: 12.02% Indels: 21
DB: 9 Gaps: 4

US-10-021-323-13 (1-609) x US-10-157-031-295 (1-146)

QY 553 GGTAAGATGTCCTCCCTTATGTAAGACGACTTGAACCGTATTCGAGAGCTGCAAG 494
DB 3 GlnGluLeuThrProGlnGlnGlnAlaGlnIleLysLysAlaPheSerAlaValAspThr 22
QY 493 AATGAGATGCTGCTGATGCTGAGAGAGCTGAATTCGTTGCTCCAGAGATTCGCT 434
DB 23 AspGlyAsnGlyThrIleAsnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 42
QY 433 GTCCATTCAGCTTGAAGAAATGAGACCCCTTATGTCGGAAGAAACATGTTGAATTCGAT 374
DB 43 ---AsnLeuSerGlnAlaGlnIleuArgLysIleSerGlnValaAspSerAspLys 61
QY 373 GAATTCGTTTATGATTCATTCATTCGAAACCA---CTGCAATGCTGATGAGCA 317
DB 62 GlnGluIleSerPheGlnGlnPheLeuThrAlaAlaArgLysAlaArgAlaGlnGln 81
QY 316 GAGAGAGAGAGAAATTCGTCATTCACGGCGCGGTGAAGAAGAAAGACATGACCTTCG 257
DB 82 -----AspLeuGln 84
QY 256 AAGCTTTTAAAGTGTTCGATTCGATGAGGATGGGATGGGAGATGAGAGCTGGA 197
DB 85 ValAlaPheArgAlaPheAspGlnAspGlyAspGlyHisIleThrValaAspGlnLeuArg 104

QY 196 TAGTCTGGAAGACTGGGTATGTAGAGTGAATAATGCGAAAAGACTCCAGACATG 137
 Db 105 ArgAlaMetAlaGlyLeuGlyGln-----ProLeuProGlnGlnGlnLeuAspAlaMet 122
 QY 136 ATTGTGATTACGACACCAATTGACAGCGCATGTTGATTTCAGAAATTCAAAACATG 77
 Db 123 TLeArgGlnAlaAspValAspGlnAspGlyArgValAsnTyrGlnGlnPheAlaArgMet 142
 QY 76 ATG 74
 Db 143 Leu 143

RESULT 15

US-09-854-731-17
 ; Sequence 17, Application US/09854731
 ; Patent No. US20020120949A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Lee, Jian Ming
 ; TITLE OF INVENTION: Plant Protein Kinases
 ; FILE REFERENCE: BB-1171
 ; CURRENT APPLICATION NUMBER: US/09/854,731
 ; CURRENT FILING DATE: 2001-05-14
 ; PRIOR APPLICATION NUMBER: 60/092,438
 ; PRIOR FILING DATE: July 10, 1998
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 17
 ; LENGTH: 639
 ; TYPE: PRF
 ; ORGANISM: Zea mays
 US-09-854-731-17

Alignment Scores:

Pred. No.: 0.00149 Length: 639
 Score: 118.00 Matches: 44
 Percent Similarity: 45.00% Conservative: 28
 Best Local Similarity: 27.50% Mismatches: 80
 Query Match: 11.12% Indels: 8
 DB: 10 Gaps: 4

US-10-021-323-13 (1-609) X US-09-854-731-17 (1-639)

QY 547 ATGTCCCTTGTAGTACGACGCTTGTATTCGAGAACTCGAAGAAATGGA 488
 Db 436 MetAsnLysLeuLysMetAlaLeuArgValIleAlaGlnLeuSerGlu----- 453
 487 GATGGCTTCTTGTAGTGTGAGAGAGCTGAATTGTTGCTCCAGAAATCGGCTCTCCAA 428
 Db 454 AspGlnIleAlaGlyLeuArgGlnMetPheLysMetIleAspAlaAspAsnSerGlyGln 473
 427 TTGAGCCTTGAAGAAATGAGCCCTTGTAGTGGAGAAACCATGTTGAATTCGATGATTC 368
 Db 474 IleThrPheGlnGlnLeuLysValGlyLeuGlnLysValGlyAlaAsnLeuGlnLysSer 493
 367 TTGTCTTTATGATTCATCTCGAACCCATCGGACATGATGATG-----GACGAGAG 314
 QY 494 GlnIleTyrAlaLeuMetGlnAlaAlaAspValAspAsnGlnIleAspTyrGly 513
 Db 313 GAGAGAGAGAAATGTCATTCACGGCGCGGTGAGAGAGAGACAGTACCTTGGCAG 254
 514 GluPheIleAlaIleThrLeuHisLeuAsnLysValGlnArgGlnAspHisLeuPheAla 533
 QY 253 GCTTTAAAGTGTGCTGAAATGAGGATGGGTGGGGGATGTGAGAGCTTGATAC 194
 Db 534 AlaPheGlnTyrPheAspLysAspGlySerGlyTyrIleThrAlaAspGlnLeuGlnVal 553
 QY 193 GTCTGGAGAGACTGGGTATGTAGTGAATAATGAGAAAGCTGACAGAGCATGATT 134
 Db 554 AlaCysGlnGlnPheGlyLeu---GlyAspValGlnLeuGlnAsp-----LeuIle 569
 QY 133 TGGTATTACAGACCAATTCAGACGGCATGTTGATTTTCAGAAATTCAAAACATGATG 74

Db 570 GlyGlnValAspGlnAspAsnAspGlyArgGlnLeuAspTyrAsnGlnPheValAlaMetMet 589
 Search completed: June 24, 2003, 14:07:46
 Job time : 32.5 sec

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 24, 2003, 13:59:57 ; Search time 25.5 Seconds
(without alignments)
4591.832 Million cell updates/sec

Title: US-10-021-323-13
Perfect score: 1100
Sequence: 1 ggaatgaatcaaatctt.....tggaactnacaatcaagg 609

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 283224 seqs, 9613422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ n2p.model -DEV=xip
-O=/cgn2.1/USPTO.spool/US10021323/runat.24062003.102235.20509/app.query.fasta_1.775
-DB=PIR_73 -QMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=Diosume2 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=psco -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10021323 @CGN.1.1.62 @runat.24062003.102235.20509 -NCPU=6 -ICPU=3
-NO_MMP -LARGEJOURN -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	312	29.4	142	2	D86348 F24J8.15 protein (
C 2	275.5	26.0	191	2	T10620 probable calcium-b
C 3	245.5	23.1	150	2	H86194 hypothetical prote
C 4	241.5	22.8	169	2	D84864 probable calcium b
C 5	239.5	22.6	205	2	S46233 allergen Bet v III
C 6	237.5	22.4	154	2	G85041 probable calmoduli
C 7	234	22.1	152	2	T06644 calmodulin homolo
C 8	216.5	20.4	195	2	T49012 calmodulin-like pr
C 9	207.5	19.6	199	2	T07365 probable calcium-b
C 10	204	19.2	145	2	T16039 hypothetical prote
C 11	199	18.8	163	2	D96753 protein calmodulin
C 12	192	18.1	157	2	D96689 calmodulin-related
C 13	191	18.0	186	2	F86379 protein F2109.28 (
C 14	184.5	17.4	146	2	F84777 probable calmoduli

C 15	178	16.8	170	2	A86317 protein T10022.19
C 16	173.5	16.4	187	2	A84532 probable calmoduli
C 17	171	16.1	149	1	MCCKA calmodulin - Yeast
C 18	169	15.9	146	2	TS1473 calmodulin-like pr
C 19	164	15.5	164	2	T10726 calmodulin - green
C 20	163	15.4	159	2	F96794 probable calmoduli
C 21	160.5	15.1	150	2	T06437 calmodulin - garde
C 22	159	15.0	149	1	MC20F calmodulin - malar
C 23	158	14.9	575	2	C85059 probable calcium d
C 24	157	14.8	149	2	UC1309 calmodulin - Stylo
C 25	155.5	14.7	160	1	TPRBCS calmodulin - fast s
C 26	155	14.6	149	1	MCTE calmodulin - Tetra
C 27	155	14.6	149	2	S28954 calmodulin - Tetra
C 28	155	14.6	150	2	T08585 calmodulin - soybe
C 29	155	14.6	151	2	UN0722 calmodulin - Pneum
C 30	154.5	14.6	160	2	A38383 troponin C, fast s
C 31	154	14.5	148	1	MCEG calmodulin - Eugle
C 32	153.5	14.5	163	1	TPCHCS troponin C, skelet
C 33	152.5	14.4	160	1	TPHUCS troponin C, fast s
C 34	152.5	14.4	162	1	TPPGCS troponin C, skelet
C 35	152	14.3	148	1	MCSW calmodulin - scall
C 36	152	14.3	148	1	MCKXM calmodulin - sea a
C 37	152	14.3	149	1	MCPM calmodulin - Param
C 38	152	14.3	150	2	T15596 troponin C, skelet
C 39	151.5	14.3	159	1	TPPGCS troponin C, skelet
C 40	151.5	14.3	163	2	JW0063 fast skeletal trop
C 41	151.5	14.3	205	2	T08398 calmodulin homolo
C 42	151	14.2	149	1	MCKHM calmodulin, striat
C 43	151	14.2	149	1	MCHUNB calmodulin-related
C 44	151	14.2	163	1	MCKM calmodulin - Chlam
C 45	150.5	14.2	163	2	JW0062 fast skeletal trop

ALIGNMENTS

RESULT 1

D86348 F24J8.15 protein (Atg121550/F24J8.7) - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-bar cross)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #ext_change 02-Nov-2001

C/Accession: D86348

R/Modelogs: A.; Eckert, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, J.
ansen, N.F.; Hughes, B.; Hutzar, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maltli, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallor
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; PMID:21016719; PMID:11130712

A/Accession: D86348

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-142 <STO>

A/Cross-references: GB:AE005172; NID:99454582; PIDN:AAF87905.1; GSPDB:GN00141

C/Genetics:

A/Map position: 1

Alignment Scores:
Pred. No.: 3.86e-21 Length: 142
Score: 312.00 Matches: 68
Percent Similarity: 63.33% Conservative: 27
Best Local Similarity: 45.33% Mismatches: 45
Query Match: 29.41% Indels: 10
DB: 2 Gaps: 3

US-10-021-323-13 (1-609) x D86348 (1-142)

QY 514 GTATTCGAGAGCTCCACAGATGAGATGAGCTGCTGTCGAGAGAGCTGAATTGG 455
DB 1 Metphenylthiureapylpsaenglnaspglvleuvalthrleuaspgluleuleutrp 20

A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-157 <STO>
 A/Cross-references: GB:AE005173; NID:g10092502; PIDN:AAG12902.1; GSPDB:GN00141
 C/Genetics:
 A/Gene: F86379
 A/Map position: 1
 C/Superfamily: calmodulin; calmodulin repeat homology
 C/Keywords: EF hand

Alignment Scores:
 Pred. No.: 6.86e-10 Length: 157
 Score: 192.00 Matches: 50
 Percent Similarity: 52.17% Conservative: 34
 Best Local Similarity: 31.06% Mismatches: 51
 Query Match: 18.10% Indels: 26
 DB: Gaps: 5

US-10-021-323-13 (1-609) x D96689 (1-157)

```

535 AGTAAGACCGACTTGAACCGCTATTCGAGAGCTCGACAGATGAGATGCTTGGT 476
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12 SerMetGluAspIleValPheGlnPheAspIleValSerMetGluAspIle 31
Qy 475 AGTCTGAGAGAGCTGAATGGTGGTCTCCAGAAATCGGGTCTGCAATTCAGCTTGA 416
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32 SerIleAspIleValSerMetGluAspIleValSerMetGluAspIleVal 50
Qy 415 GAATTCGAGCCCTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 374
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 51 GluThrIleValMetGluValPheAspIleValSerMetGluAspIleVal 70
Qy 373 GAATTCGCTTCTTATGATTCATTCGACACCTGCGACATGCTGCTGCTGCTG 314
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 71 GluPheValAlaLeuPheGlnIle-----SerAspGlnSer 82
Qy 313 GAGGAGAGAGAAATGCTCATTACGCGCGCGGTGAAGAGACAGTACCTTGGCAG 254
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 83 SerAsnAsnSerAlaIle-----ArgAspLeuIleVal 93
Qy 253 GCTTTTAAAGTGTGACTTGAATGGGATGGGATGGGATGGGATGGGATGGGAT 194
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 94 AlaPheAspLeuIleValSerMetGluAspIleValSerMetGluAspIleVal 113
Qy 193 GTGCTGGGAGAGAGCTGATGAGTGAATATGTAAGAACTGCGACGATGAT 134
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 114 ValMetGluAsnLeuGly-----GluIleCysSerIleGlnAspCysGlnArgMetIle 131
Qy 133 TGGTATTATCGACACCAATTCAGACGCGCATGGTGTGATTTTCAAGATTCAAAACATGATG 74
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 132 AsnIleValAspSerAspGlyAspGlyCysValAlaPheGlnGluPheIleMetMet 151
Qy 73 TTA 71
    |||
Db 152 Met 152

```

RESULT 13

F86379 protein F21J9.28 [imported] - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C/Accession: F86379
 A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A/Reference number: A86141; MUID:21016719; PMID:11130712
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-157 <STO>
 A/Cross-references: GB:AE005173; NID:g10092502; PIDN:AAG12902.1; GSPDB:GN00141
 C/Genetics:
 A/Gene: F86379
 A/Map position: 1
 C/Superfamily: calmodulin; calmodulin repeat homology
 C/Keywords: EF hand

A/Accession: F86379
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-186 <STO>
 A/Cross-references: GB:AE005172; NID:g9743349; PIDN:AAF97973.1; GSPDB:GN00141
 C/Genetics:
 A/Gene: F21J9.28
 A/Map position: 1
 C/Superfamily: calmodulin; calmodulin repeat homology
 C/Keywords: EF hand

Alignment Scores:
 Pred. No.: 8.82e-10 Length: 186
 Score: 191.00 Matches: 54
 Percent Similarity: 49.68% Conservative: 24
 Best Local Similarity: 34.39% Mismatches: 51
 Query Match: 18.00% Indels: 28
 DB: Gaps: 5

US-10-021-323-13 (1-609) x F86379 (1-186)

```

Qy 526 GACTTCGACGCGCTATTCGAGAGCTCGACAGATGAGATGCTTGGTACGAG 467
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 37 GluLeuGluAlaValPheIleValSerMetGluAspIleValSerMetGlu 56
Qy 466 GAGCTGAATGCTGCTCCAGAGATCGGCTGCTGCTGCTGCTGCTGCTGCTG 407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 57 GluLeuGlyAlaIleMetThrSerLeuGlyHis---GluValProGluGluIleVal 75
Qy 406 CCCTTAGTG-----GGAAAACATGTTGAATTCGATGATTCCTTG 365
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76 LysAlaIleThrGluIleAspArgIleValSerMetGluAspIleValSerMet 95
Qy 364 TCTTTTATGATTCATTCCTGAAACCCATGCGACATGCTGCTGCTGCTGCTG 305
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 96 GluLeuAsn-----ThrIleGlyMetAspGlnIleValSerMetGlu 108
Qy 304 GAATTCGCTATTCACGCGCGCGGTGAAGAGACAGTACCTTCGAAAGCTTTAAA 245
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 109 Glu-----AsnLeuIleValSerMetGluAspIleValSerMetGlu 116
Qy 244 GTGCTTGAATTCGAGAGAGTGGTGGGAGATGAGAGAGCTTGAATACGCTTGG 185
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 117 ValIleAspIleAspGlyValSerIleSerAlaGluIleValSerMetGlu 136
Qy 184 AGACTGGTATGAGAGTGAATATGCTGAAAAGATCTGCGAGCATGATTTGGTAT 125
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 137 SerLeuIle-----AspGluCysSerIleAlaGluValArgIleValSer 154
Qy 124 GACCAATTCAGACGCGCATGCTGATTTCAAGATTCAAAACATGATG 74
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 155 AspIleAspGlyValSerMetGluIleValSerMetGluIleValSerMet 171

```

RESULT 14

F84777 probable calmodulin [imported] - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Mar-2001
 C/Accession: F84777
 A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A/Reference number: A84420; MUID:20083487; PMID:10617197
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-146 <STO>
 A/Cross-references: GB:AE002093; NID:g4510338; PIDN:AAD21447.1; GSPDB:GN00139
 C/Genetics:
 A/Gene: At2g36180
 A/Map position: 2
 C/Superfamily: calmodulin; calmodulin repeat homology
 C/Keywords: EF hand

Alignment Scores:
 Pred. No.: 3.4e-09 Length: 146
 Score: 184.50 Matches: 51
 Percent Similarity: 53.50% Conservative: 33
 Best Local Similarity: 32.48% Mismatches: 26
 Query Match: 17.39% Indels: 47
 DB: 2 Gaps: 6

US-10-021-323-13 (1-609) x F84777 (1-146)

QY 523 TTGCAAGCCGTTATCGAAGCTCGACAGATGAGATGGCTTCGTTAGTCGTGAGAG 464
 Db 3 MetlaagluilepnegluserValaspLyasnlYaspgLyLleLeuTrpaspLiu 22
 QY 463 CTGAATGGCTTCCTCCAGATCGGCTCTGCTCAATTCAGCCTTGAAGATGAGAGCC 404
 Db 23 Phehlaglualelle---ArgValPheSerProGlnIleThrsenrIugluileasp--- 40
 QY 403 TTACTGGGAAACCATGTTGAATCTTGATGAATCTTGTCTTTATGAATCCATCTCG 344
 Db 41 -----LysMetPheIleValleuasp----- 47
 QY 343 AACCACTGCAKATGTCGTGACGAGAGAGAGAGAGAA-----TTGGTC 296
 Db 48 -----ValaspGlyaspGlyGlnIleaspaspValGluPheAlaSerCysLeuMet 64
 QY 295 ATTCACGCGCGCGGTGA-----GAAGAGACAGTGAACCTTGGCAGAGCTTTTAA 245
 Db 65 ValasnIyglYglYglYasprThrdIugluIuValIleMetValgluIlePheasp 84
 QY 244 GTGTTGACTTGAATGGGATGGGTGGGGGATGTGAGAGCTTGAATACGTGCTGGA 185
 Db 85 LeuTrpaspMetaspGlyaspGlyLysIleSerAlaSerGluIleHisValIleuLys 104
 QY 184 AGACTGGGTATGTAGAGTGAATAATGCGAAAGAAAGCTGCGAGAGCATGATTTGGTATTAC 125
 Db 105 ArgIeuGlyMet-----GluYasHleThrMetGluaspCysValIleMetValGlnThrVal 122
 QY 124 GACCACTATTCAGACGCGATGTTGATTTCAAGATTCAGAAACATGATG 74
 Db 123 AspLyaspSeraspGlyaspPheValasnPhegluIuPheLysIleMetMet 139

RESULT 15

A86317

protein T10022.19 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Apr-2001

Accession: A86317

..Theologis, A.; Becker, J.R.; Palm, C.J.; Federpiet, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Hutzar, L.
 Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzalli,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 Kerr, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.M.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: A86317

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-170 <STO>

A/Cross-references: GB:AE005172; NID:98671778; PIDN:AAF8384.1; GSPDB:GN00141

C/Genetics:

A/Map position: 1

C/Superfamily: calmodulin repeat homology

C/Keywords: EF hand

Alignment Scores:

Pred. No.: 1.43e-08

Length:

170

Score: 178.00 Matches: 49
 Percent Similarity: 49.38% Conservative: 30
 Best Local Similarity: 30.62% Mismatches: 49
 Query Match: 16.78% Indels: 32
 DB: 2 Gaps: 4

US-10-021-323-13 (1-609) x A86317 (1-170)

QY 526 GACTTGCAAGCCGTTATTCGAGAGCTCGACAGATGAGATGGCTTCGTTAGTCGTGAG 467
 Db 23 GluIeuYasLyasValPheaspGlnPheaspSeraspGlnIleaspGlyLysIleSerValIleu 42
 QY 466 GAGCTGAATGGCTTCCTCCAGATCGGCTCTGCTCAATTCAGCCTTGAAGATGAGAG 407
 Db 43 GluIeuGlyIyValPheLysAlaMetGlyThr---SerYrThrGluThrGluLeuasn 61
 QY 406 CCCTTAGTGGGAAA-----CCAGTTGAACCTGATGAATCTTG 365
 Db 62 ArgValIleuGluIuValaspThraspArgaspGlyTyrlleasnleuaspGluPheSer 81
 QY 364 TTCTTTATGAATCCATCTCGAAACCACTGCGACATGTGTGTGACGAGAGAGAGAG 305
 Db 82 ThrIeuYasraserSerSerSerAla----- 90
 QY 304 GAATGGTCAATTCACGCGCGGTGAAGAAAGACAGTGAACCTTGGAGAGCTTTTAA 245
 Db 91 -----AlaGluIleArgaspAlaPheasp 98
 QY 244 GTGTTGACTTGAATGGGATGGGTGGGGGATGTGAGAGAGCTTGAATACGTGCTGGA 185
 Db 99 LeuTrpaspGlnaspLyasnglyLeuIleSerAlaSerGluIleuHisGlnValIleuasn 118
 QY 184 AGACTGGGTATGTAGAGTGAATAATGCGAAAGAAAGCTGCGAGAGCATGATTTGGTATTAC 125
 Db 119 ArgIeuGlyMet-----SerCysSerValGluaspCysThrArgMetIleGlyProval 136
 QY 124 GACCACTATTCAGACGCGATGTTGATTTCAAGATTCAGAAACATGATGATGATTC 65
 Db 137 AspLaspaspGlyaspGlyaspValasnPhegluIuPhegluIuYasMetMetThrSerSer 156

Search completed: June 24, 2003, 14:06:04
 Job time : 28.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_nzp model

Run on: June 24, 2003, 14:03:28 ; Search time 13 Seconds
(without alignments)
3886.013 Million cell updates/sec

Title: US-10-021-323-13
Perfect score: 1100
Sequence: 1 ggcgaatgaatcaactttc.....tgaagctncaaatcaagg 609

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Arched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 225764

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+npz.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool/US10021323/runat.24062003.102234.20466/app.query.fasta_1.775
-DB=SwissProt.40 -QFMT=fastan -SUFFIX=isp -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -STAR=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLGN=200 -THR_SCORE=perc -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US10021323 @CGN 1.1 26 @runat.24062003.102234.20466 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOUTRY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	239.5	22.6	205	ALB3_BETVE	P43187 betula verr
C 2	218	20.5	161	TCH2_ARATH	P25070 arabidopsis
C 3	207.5	19.6	199	CAST_SOLTU	O09011 solanum tub
C 4	185.5	17.5	171	ALB8_OLEEU	O96770 olea europa
C 5	175	16.5	149	CALM_CANAL	P23286 candida alb
C 6	159	15.0	149	CALM_PLARA	P24044 plasmodium
C 7	157	14.8	148	CALM_STYLE	P27166 stylyonchia
C 8	155.5	14.7	159	TPCS_RABIT	P02588 corydalis
C 9	155	14.6	148	CALM_TERYP	P02598 tetrahyena
C 10	155	14.6	151	CALM_PNECA	P41041 pneumocysti
C 11	154.5	14.6	159	TPCS_MOUSE	P20801 mus muscula
C 12	154	14.5	148	CALM_EUGER	P11118 euglena gry
C 13	153.5	14.5	162	TPCS_CHICK	P02588 gallus gall
C 14	152.5	14.4	159	TPCS_HUMAN	P02589 homo sapien
C 15	152.5	14.4	162	TPCS_RANES	P02589 rana esculie
C 16	152	14.3	148	CALM_METSE	P02596 metridium s
C 17	152	14.3	148	CALM_PATRE	P07463 parametium s
C 18	152	14.3	148	CALM_PATSP	P02595 patinopecte

C 19	151.5	14.3	159	1	TPCS_PIG	P02587 sus scrofa
C 20	151	14.2	148	1	CALM_HUMAN	P27482 homo sapien
C 21	151	14.2	148	1	CALS_CHICK	P02597 gallus gall
C 22	151	14.2	162	1	CALM_CHIRE	P04352 chlamydomon
C 23	150	14.1	193	1	E631_DROME	P48593 drosophila
C 24	149.5	14.1	162	1	TPCS_MEIGA	P10246 meleagris g
C 25	149	14.0	148	1	CALM_BLAEM	O9f56 blasocycladi
C 26	148.5	14.0	173	1	CALM_BOVIN	O9n498 bos taurus
C 27	148	13.9	148	1	CALM_FLEOS	O94739 pleurotus o
C 28	148	13.9	162	1	CAB2_BOVIN	O9n149 bos taurus
C 29	148	13.9	165	1	POC2_JUNOX	O64943 juniperus o
C 30	147.5	13.9	173	1	CAB5_HUMAN	O9n866 homo sapien
C 31	147	13.9	148	1	CALM_PHYIN	P21265 phytophor
C 32	147	13.9	148	1	CALM_PLECO	P11120 pleurotus c
C 33	147	13.9	148	1	CALM_PYUSP	P11121 pyridae sp
C 34	147	13.9	148	1	CALM_TRYCR	P18061 trypanosoma
C 35	146.5	13.8	148	1	CLM4_MOUSE	O9jme3 mus musculu
C 36	146.5	13.8	161	1	CALM_CABEL	P04630 caenorhabdi
C 37	146	13.8	148	1	CALM_TRYBB	P13668 solanum tub
C 38	145	13.7	148	1	CALM_SOLTU	O9j1k4 mus musculu
C 39	145	13.7	215	1	CAB2_MOUSE	O9j1k3 mus musculu
C 40	144.5	13.6	173	1	CAB5_MOUSE	P30188 arabidopsis
C 41	144.5	13.6	215	1	CALM_ARATH	P27161 lycopersico
C 42	144	13.6	148	1	CALM_LYCES	P04464 triticum ae
C 43	144	13.6	149	1	CALM_WHEAT	P27164 petunia hyb
C 44	144	13.6	193	1	CAL2_PETRY	O9npd3 homo sapien
C 45	144	13.6	219	1	CAB2_HUMAN	

ALIGNMENTS

RESULT 1
ID ALB3_BETVE STANDARD; PRT; 205 AA.
AC P43187;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcium-binding allergen Bet v 3 (Bet v III).
OS Betula verrucosa (White birch) (Betula pendula).
GN Eukaryota; Viridiplantae; Magnoliophyta; Embryophyta; Tracheophyta;
OC Eukaryota; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Euphorbia; Euphorbiaceae; Betula.
OX NCBI_TaxID=3505;
RN [1]
RP SEQUENCE FROM N.A., AND SYNTHESIS OF CALCIUM-BINDING SITES.
RC TISSUE=Pollen;
RX MEDLINE=94341260; PubMed=7520389;
RA Seiberter S., Scheiner O., Kraft D., Lonsdale D., Valenta R.;
RT "Characterization of a birch pollen allergen, Bet v III, representing
RT a novel class of Ca2+ binding proteins: specific expression in mature
RT pollen and dependence of patients' IGE binding on protein-bound
RT Ca2+";
RL EMO J. 13:3481-3486(1994).
CC -!- FUNCTION: COULD BE INVOLVED IN CALCIUM METABOLISM IN POLLEN. BINDS
CC -!- 3 CALCIUM IONS.
CC -!- SIMILARITY: CONTAINS 3 EF-HAND CALCIUM-BINDING DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC EMBL; X12567; CA55854.1; -
CC HSSP; P02593; 1CDM.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; efhand; 3.
DR ProDom; PD000012; EF-hand; 2.
DR SMART; SM00054; Efh; 3.

DR PROSITE; PS00018; EF HAND; 3.
 KW Allergen; Calcium-binding; Repeat.
 FT CA BIND 49 EF-HAND 1 (POTENTIAL).
 FT DOMAIN 85 ANCESTRAL CALCIUM SITE 2.
 FT CA BIND 143 154 EF-HAND 3 (POTENTIAL).
 FT CA BIND 181 192 EF-HAND 4 (POTENTIAL).
 FT SEQUENCE 205 AA, 23141 MM, 350B386C0FE05C CRC64;
 SQ
 Alignment Scores:
 Pred. No.: 2,45e-13 Length: 205
 Score: 239.50 Matches: 63
 Percent Similarity: 57.93% Conservative: 32
 Best Local Similarity: 38.41% Mismatches: 58
 Query Match: 22.57% Gaps: 11
 Gaps: 6
 US-10-021-323-13 (1-609) x ALL3_BERVE (1-205)
 QY 538 CTTAGTAAGACCGCTTGCACACCGCTATTCAGAGAGCTCAGAGATGAGATGCTTC 479
 36 LENAANTHRLAARGARGARGILEPHEASPLEUPHEAPLYASNBERSAPGLY11e 55
 QY 478 GTTAGTCGAGAGAGCTGGAATGCTGCTCCAGAGAGATCGGCTGTCGCAATTCAGCCTT 419
 56 ILEHVALASPGLULSERATRGALALEUASNLLEUENGLY--LEUGLUTHPASPLEU 74
 QY 418 GAAAGATTCGAGCCCTTAGTG------GGAAACGATGTTGAACCTTG 377
 75 SERGLULUENGLUSERTHVALYSERPHERHARGLUGLUGLIANILEGLYLEUGLPH 94
 QY 376 GATCAATTCCTGTTCTTTTGAATTCATCTCGAACCA--CTGGCAGCATGGTGT-- 323
 95 GLUASPHEHLESERLEUHLGLNLERLEUASNPSERTYRPHALATYRGLYGLU 114
 QY 322 GAGCAAGAGAGAGAGAGATTCATTCACGCCGCGGTGAAGAAAGAGAGAGAC 263
 115 ASPGLUASPAPANGULGLUASPMEFARGLYSERILEUENGLUGLUGLUALASP 134
 QY 262 CTTGCGAAGGCTTTAAAGTGTTCATTCATTCAGAGATGAGATGAGATGAGAG 203
 135 SERPHEGLYGLYPHELYVALPHEASPLUASPLGLYSPGLYRLLSEALARGGLU 154
 QY 202 CTTGAATACGCTGCTGGAGAGACTGGATGTGAGCTGAAATAGTGAAGAAAGAC--TGC 146
 155 LEUGLMEVALLEUGLYLEUGLYPHE--SERGLUGLYSERGLULLEASPARGLY 173
 QY 145 AGGAGCATGTTTGTGATTCAGACACCAATTCAGAGAGATGAGATTCAGATTC 86
 174 GLULYSMEGLLEVALSERVALASPSEFARNARGAPGLYARGVALASPHEPHEGLUPHE 193
 QY 85 AAAACATGATG 74
 194 LYSASPHEMET 197
 DB
 RESULT 2
 TCH2_ARATH STANDARD; PRT; 161 AA.
 ID TCH2_ARATH
 AC P25070; O22592;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 15-JUN-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Calmodulin-related protein 2, touch-induced.
 GN TCH2 OR AT5G37770 OR K22F20.1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxId=3702;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Johnson K.A., Braam J.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=99087489; PubMed=9872454;
 RA Nakamura Y., Sato S., Asamizu E., Kaneko T., Kotani H., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VII.
 RT Sequence features of the regions of 1,013,767 bp covered by sixteen
 RT physically assigned P1 and TAC clones";
 RL DNA Res. 5:297-308(1998).
 RN [3]
 RP SEQUENCE OF 26-70 FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=90150263; PubMed=2302732;
 RA Braam J., Davis R.W.;
 RT "Rain-, wind-, and touch-induced expression of calmodulin and
 RT calmodulin-related genes in Arabidopsis";
 RL Cell 60:357-364(1990).
 RN [4]
 RP 3D-STRUCTURE MODELING OF 7-158.
 RX MEDLINE=97189489; PubMed=9037719;
 RA Khan A.R., Johnson K.A., Braam J., James M.N.G.;
 RT "Comparative modeling of the three-dimensional structure of the
 RT calmodulin-related TCH2 protein from Arabidopsis";
 RL Proteins 27:144-153(1997).
 CC
 CC -1- SIMILARITY: CONTAINS 4 EF-HAND CALCIUM-BINDING DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sb-sib.ch).
 CC
 CC EMBL; AF026473; AAB82713.1; -;
 CC EMBL; AB016873; BAB10353.1; -;
 CC PIR; C34669; C34669.
 CC PDB; 1AVJ; 28-JAN-98.
 CC InterPro: IPR002048; EF-hand.
 CC Pfam; PF00036; efhand; 4.
 CC ProDom; PD000012; EF-hand; 2.
 CC SMART; SM00054; Efn; 4.
 CC PROSITE; PS00018; EF HAND; 4.
 KW Calcium-binding; Repeat; 3d-structure.
 FT CA BIND 26 37 EF-HAND 1 (POTENTIAL).
 FT CA BIND 62 73 EF-HAND 2 (POTENTIAL).
 FT CA BIND 103 114 EF-HAND 3 (POTENTIAL).
 FT CA BIND 139 150 EF-HAND 4 (POTENTIAL).
 FT CONFLICT 54 54 T -> Y (IN REF. 3).
 SQ SEQUENCE 161 AA; 17546 MM; A7324A5C0E0CBDB9 CRC64;
 Alignment Scores:
 Pred. No.: 1.78e-11 Length: 161
 Score: 218.00 Matches: 55
 Percent Similarity: 52.41% Conservative: 32
 Best Local Similarity: 33.13% Mismatches: 43
 Query Match: 20.55% Indels: 36
 Gaps: 6
 DB: 1
 US-10-021-323-13 (1-609) x TCH2_ARATH (1-161)
 QY 535 AGTAAGACCGCTTGCACACCGCTATTCAGAGAGCTCAGAGATGAGATGCTTCGT 476
 14 SERMECAPSPHELYSPVALPHEGNARGPHEASPLUASNPGLYSPGLY11e 33
 QY 475 AGTCGAGAGAGCTGGAATGCTGCTCCAGAGAGATCGGCTGCTC----- 431
 34 SERVALASPGLULUENGLUGLULIARGALALEUASERPHTHRLASERPROGLUGLU 53
 QY 430 -----CATTCAGCTTGAAGATTCGAGCCCTTAGCGGAAACCA 389

```

RESULT 3
CAST_SOLTU STANDARD; PRT; 199 AA.
ID CAST_SOLTU
AC Q09011;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Calcium-binding protein CAST.
OS Solanum tuberosum (Potato).
OS Eubacteria; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;
OC Aspergillales; Ascomycota; Ascomycota; Ascomycota; Ascomycota;
OX NCBI_Taxid=4113;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Kennebec; TISSUE=tuber;
RX MEDLINE=94143493; PubMed=8310073;
RA Gellatly K.S., Lefevre D.D.;
RT "Identification of a cDNA clone coding for a novel calcium-binding
RL protein from potato tuber";
RL Plant Physiol. 101:1405-1406(1993).
C1 - FUNCTION: NOT KNOWN. PROBABLY BINDS 3 CALCIUM IONS.
C1 - SIMILARITY: CONTAINS 3 EF-HAND CALCIUM-BINDING DOMAINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL, L02830; AAA33811.1; -.
DR HSPF, P02593; ICTR.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 3.
DR ProDom; PD000012; EF-hand; 2.
DR SMART; SM00054; EFh; 3.
DR PROSITE; PS00018; EF_HAND; 1.
KM Calcium-binding; Repeat.
FT DOMAIN 24 28
FT CA BIND 49 60 POLY-SER.
FT DOMAIN 85 96 EF-HAND 1 (POTENTIAL).
FT CA BIND 138 149 ANCESTRAL CALCIUM SITE 2.
FT CA BIND 176 187 EF-HAND 3 (POTENTIAL).
FT CA BIND 199 226 EF-HAND 4 (POTENTIAL).
SQ SEQUENCE 199 AA; 22606 MW; D8F7C3DDEE5415A2 CR664;

```

US-10-021-323-13 (1-609) x CAST_SOLTV (1-199)

OY		583	GAGTTTAAAGTGTCATCATATC-----CTCTTTGGGTGTAATAAAGTCCCCCCTTAGT	533
Dd		6	AspGluAsnHisLeuAspGluPheSerGlnIleuValLeuGlyLeuValProSerSer	25
OY		532	AAGACCGAC-----TTGCACCGCGTATT	509
Dd		26	SerSerSerPheArgLeuArgSerProSerLeuAsnSerIleArgLeuArgAlaArgIlePhe	45
OY		508	GAGAAGCTGCACAAGATGGAATAGGCTTCGTACTCTGGAGAGACCTGAATTGGTTGCTC	449
Dd		46	AspValAPheAspArgAsnHisAspCysLeuIleSerValGIuGIuLeuSerGlnAlaLeu	65
OY		448	CAGAGATCGGGGTCTCCCAATTCAAGCTTGAAAGATGGAGCCCTAGTGGGA-----	395
Dd		66	AsnLeuLeuGly---LeuAspAlaAspLeuSerGlnIleGluSerMetValLysLeuHis	84
OY		394	---AAACCA-----TGTTGAACCTTGATGAATCTTGTTCTTTATGAATCCATC	347
Dd		85	IleLysPheGlnAsnThrGlyLeuArgPheGluAsnProGluIuThrLeuHisArgSerLeu	104
OY		346	TGGAACCCACTGGCACATGTGTGTGACGAAGAAGAGAGAGGAATTGGTCATTCACGCC	287
Dd		105	---AsnAspValAPhePheGly-----SerLysCysGlnAspLysLeuGlyLeuAsnPro	121
OY		286	GGCGGGTGAAGAAGAACACATGACCTTGACGATTTTAAAGGTTTGAACCTTGAAAGG	227
Dd		122	AspProAlaGlnAspGluSerAspLeuLysGlnValAlaPheAspValAPheAspGluAsnGly	141
OY		226	GATGGGTGGGGGATGTAGAGAGGCTGGAATACGTGCTGGAGAGACTGGGTATGTAGAGT	167
Dd		142	AspGlyPheIleSerAlaLysGluLeuGlnValValLeuGluLysLeuGlyLeu---Pro	160
OY		166	GAAAATATGTGAAAAAGAC---TGACGAGACATGATTTGGTATTCACACCAATTGAC	110
Dd		161	GIuGIySerGluIleAspArgValGIuMetMetIleSerSerValGIuGlnAsnHisAsp	180
OY		109	GGCATGGTGTATTTCAAAATTCAAAAACATGATG	74
Dd		181	GIyArgValAspPheGlnPheLysAspMetMet	192
<hr/>				
RESULT 4				
ALeU_OLeEU STANDARD; PRT; 171 AA.				
AC	Q9M7R0; Q9M7O9;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Calcium-binding allergen Ole e 8 (PcA18/PcA23).			
OS	Olea europaea (Common olive).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
OC	Asteridae; euasterids I; Lamiales; Oleaceae; Olea.			
OX	NCBI_TaxId=4146;			
RN				
RP	SEQUENCE FROM N.A.			
RC	TISSUE=PolLEN;			
RX	MEDLINE=20115491; PubMed=10648940;			
RA	Ledesma A.; Villalba M.; Rodriguez R.;			
RT	"Cloning, expression and characterization of a novel four EF-hand			
RU	Ca(2+)-binding protein from olive pollen with allergenic activity.";			
PBS	lett. 466:192-196(2000).			
-1- SIMILARITY: CONTAINS 4 EF-HAND CALCIUM-BINDING DOMAINS.				

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL AF078679; AAF31151.1; -
DR EMBL AF078680; AAF31152.1; -
DR HSSP; P02593; 1CTR.
DR InterPro; IPR003299; Calflagin.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 4.
DR PRINTS; PRO1362; CALFLAGIN.
DR ProDom; PD000012; EF-hand; 2.
DR SMART; SM00054; Eph; 4.
DR PROSITE; PS00018; EF-hand; 4.
KW Calcium-binding; Repeat; Allergen.
CA_BIND 29 40 EF-HAND 1 (POTENTIAL).
CA_BIND 65 76 EF-HAND 2 (POTENTIAL).
CA_BIND 105 116 EF-HAND 3 (POTENTIAL).
CA_BIND 141 152 EF-HAND 4 (POTENTIAL).
FT VARIANT 43 44 GV -> CA.
FT VARIANT 58 59 G -> A.
FT VARIANT 60 60 I -> M.
SQ SEQUENCE 171 AA, 18907 MW, 34C430E468E2B284 CRC64;

Alignment Scores:
Pred. No.: 1,23e-08 Length: 171
Score: 185.50 Matches: 51
Percent Similarity: 50.96% Conservative: 32
Best Local Similarity: 32.48% Mismatches: 25
Query Match: 17.48% Indels: 25
Gaps: 4

US-10-021-323-13 (1-609) x ALB_OLEU (1-171)

QY 526 GACTTGCAACGCGTATTTCGAAAGCTCGACAAAGATGAGATGCTTACTGTCGAG 467
DB 20 GlnValGlnGlnValPheAsnArgPheAspAlaAsnGlyAspGlyAsp 39
QY 466 GACGTGAATGGTGGTCCGAGAGATGCGTGTCCAAATCGCTTGAAGATGAG 407
DB 40 GlnLeuAlaGlyValLeuValAlaLeuGlySer--AsnThrSerIysGlnGlnIleGly 58
QY 406 CCCTTAGTGGGAAA-----CCATGTTGAATTGATGATTTCTTG 365
DB 59 ArgIleMetGlnGlnIleAspThrAspIysAspGlyPheIleAsnValGlnGlnPheAla 78
QY 364 TTTCTTTATGATTCATTCGAAACCACTGCGACATGCTGTGACGAAGAGAGAGAG 305
DB 79 AlaPheValIysAlaGlnThrAspProTyrProSerSerIysGlyGlnAsnGlu----- 96
QY 304 GAATGGTCATTCACGGCGCGGCGTGAAGAAGACGTACCTTGGAGAGCTTTAA 245
DB 97 -----LeuIysGlnAlaPheGln 102

QY 244 GTGTGACTTGAATGGAGATGGTGGGGGATGTGAGAGCTTGAATCGTCTGGA 185
DB 103 LeuTyrAspGlnAsnIleAsnGlyLeuIleSerSerValGlnLeuHisIleLeuThr 122
QY 184 AGACTGGGATGTGAGAGTGAATAATAGTGAAGAACTGCGAGAGCATGATTTGGTATTAC 125
DB 123 ArgGlnGly-----GluArgTyrAlaGlnIleAspGlyValAlaGlnMetIleIysSerVal 140
QY 124 GACCAATTCAGACGGCATGCTGATTTCCAGAAATTCATAAATCAATGAG 74
DB 141 AspSerAspGlyAspGlyTyrValSerPheGlnIlePheIysIysMetMet 157

RESULT 5
CALM_CANAL
ID CALM_CANAL STANDARD; PRT; 149 AA.

AC P23286;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Calmodulin (Cam).
GN CML1.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_Taxid=5476;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=92039062; PubMed=1937040;
RA Saportis S.M., Sypherd P.S.;
RT "The isolation and characterization of a calmodulin-encoding gene
RT (CML1) from the dimorphic fungus Candida albicans.";
RL Gene 106:43-49(1991).
CC -1- FUNCTION: CALMODULIN MEDIATES THE CONTROL OF A LARGE NUMBER OF
CC ENZYMES BY CA(++) AMONG THE ENZYMES TO BE STIMULATED BY THE
CC CALMODULIN-CA(++) COMPLEX ARE A NUMBER OF PROTEIN KINASES AND
CC PHOSPHATASES.
CC -1- MISCELLANEOUS: THIS PROTEIN HAS FOUR FUNCTIONAL CALCIUM-BINDING
CC SITES.
CC -1- SIMILARITY: CONTAINS 4 EF-HAND CALCIUM-BINDING DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M61128; AAA4331.2; -
DR PIR; JU0276; MCCXA.
DR HSSP; P02593; 1CDM.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 4.
DR ProDom; PD000012; EF-hand; 2.
DR SMART; SM00054; Eph; 4.
DR PROSITE; PS00018; EF-hand; 4.
KW Calcium-binding; Repeat.
FT CA_BIND 21 32 EF-HAND 1.
FT CA_BIND 57 68 EF-HAND 2.
FT CA_BIND 94 105 EF-HAND 3.
FT CA_BIND 130 141 EF-HAND 4.
SQ SEQUENCE 149 AA, 16492 MW, 2679FCF975B9B04C CRC64;

Alignment Scores:
Pred. No.: 1e-07 Length: 149
Score: 175.00 Matches: 48
Percent Similarity: 49.06% Conservative: 30
Best Local Similarity: 30.19% Mismatches: 65
Query Match: 16.49% Indels: 16
Gaps: 3

US-10-021-323-13 (1-609) x CALM_CANAL (1-149)

QY 550 AAAATGTCCTCCCTTAGTAAGACCGACTTGAACGGATTCGAGAACTGCAAGAAAT 491
DB 4 LysLeuSerGlnGlnGlnIleAlaGlnPheIysGlnAlaPheSerIleuPheAsp 23
QY 490 GGAATGCGCTTGTATGCTGAGAGAGCTGAAATGCTTCTCCAGAGAAATCGGCTCTGC 431
DB 24 SerAspGlyLysIleThrThrIysGlnLeuGlyThrValMetArgSerLeuGln---- 42
QY 430 CAATTAGCTTGAAGAATTCGAGCGCTTGAAGGAAAACATGTTGAATTCGATGA 371
DB 43 AsnProSerGlnSerIleuLeuThrAspMetIleAsnGlnValAspValAsnSerAspGly 62
QY 370 TTTCTGCTTTTATGATTCATTCGAAACCACTGCGACATGCTGTGACGAAGAGAG 311
DB 63 SerIleAspPheProGlnPheLeuThrMetMetAlaArgIysMetLysAspThrAspSer 82

QY 310 GAGAGGAATTGTCATTACACGGCGGGTGAAAGAAAGACAGTACCTTCCAGGCT 251
 DB 83 GUAAGAGT-----11leagGua1a 89
 QY 250 TTTAAAGTGTGTTGACTGTAATGGAGTGGGAGTGTGAGAGCTTGAATACGTG 191
 DB 90 PheLyValPheAspArgAsnGlyAspGlyLysLeuSer1a1aGluLeuArgHisLeu 109
 QY 190 CTGGGAAGACTGGGTATGTAGTGAATAATGTGAAAAGACTGCAGAGCATGATTGG 131
 DB 110 LeuThSer1leGly-----GluLyLeuSerAspAlaAspValAspGluMet1leLys 127
 QY 130 TATTAGACACCAATTCAGACGCGATGTTGATTGTAAGATTCAAAAATCATG 74
 DB 128 GUAAlaAspThrAsnAsnAspGlyGlu1leAsp1leGlnGlnPheThSerLeu 146
 RESULT 6
 CALM_PLAFA STANDARD; PRT; 149 AA.
 P24044;
 01-MAR-1992 (Rel. 21, Created)
 01-MAR-1992 (Rel. 22, Last sequence update)
 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Calmodulin.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=583;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Isolate FC27;
 RX MEDLINE=92008475; PubMed=1915742;
 RA Coman A.F., Galatis D.;
 RT "Plasmodium falciparum: the calmodulin gene is not amplified or
 overexpressed in chloroquine resistant or sensitive isolates";
 RL Exp. Parasitol. 73:269-275 (1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91304518; PubMed=1852174;
 RA Robson K.J.H., Jennings M.W.;
 RT "The structure of the calmodulin gene of Plasmodium falciparum";
 RL Mol. Biochem. Parasitol. 46:19-34 (1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93376007; PubMed=8366883;
 RA Robson K.J.H.;
 RT "Sequence diversity in the intron of the calmodulin gene from
 Plasmodium falciparum";
 RL Mol. Biochem. Parasitol. 60:1-8 (1993).
 CC -1- FUNCTION: CALMODULIN MEDIATES THE CONTROL OF A LARGE NUMBER OF
 ENZYMES BY CA(++) . AMONG THE ENZYMES TO BE STIMULATED BY THE
 CALMODULIN-CA(++) COMPLEX ARE A NUMBER OF PROTEIN KINASES AND
 PHOSPHATASES.
 CC -1- MISCELLANEOUS: CALMODULIN IS NOT INVOLVED IN THE MECHANISM OF
 CHLOROQUINE RESISTANCE.
 CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.1sb-sib.ch/announce/
 or send an email to license@sib-sib.ch).
 CC
 CC EMBL; MS9349; AAA29509.1; -
 CC EMBL; X56950; CAA40264.1; -
 CC EMBL; M59770; AAA29510.1; -
 CC EMBL; M99442; AAA29508.1; -
 CC PIR; S21813; S21813.
 CC PIR; B45594; B45594.
 CC PIR; A49774; A49774.
 CC HSSP; P02593; 1CDM.
 RP SEQUENCE FROM N.A.

DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; ehand; 4.
 DR ProDom; PD000012; EF-hand; 2.
 DR SMART; SM00054; EFh; 4.
 DR PROSITE; PS00018; EF_HAND; 4.
 KM Calcium-binding.
 FT CA_BIND 21 32 EF-HAND 1.
 FT CA_BIND 57 68 EF-HAND 2.
 FT CA_BIND 94 105 EF-HAND 3.
 FT CA_BIND 130 141 EF-HAND 4.
 FT CONFICT 28 30 MISSING (IN REF. 1).
 SQ SEQUENCE 149 AA; 16931 MW; 30D806FDA42BC173 CRC64;
 Alignment Scores:
 Pred. No. 2,49e-06 Length: 149
 Score: 159.00 Matches: 44
 Percent Similarity: 50.00% Conservative: 39
 Best Local Similarity: 26.51% Mismatches: 55
 Query Match: 14.99% Indels: 28
 DB: 1 Gaps: 6
 US-10-021-323-13 (1-609) x CALM_PLAFA (1-149)
 QY 550 AAAATGCCCCCTTATAGACCGACTTGCAACGCGTATTGGAAGAGTCGACAAAGAT 491
 DB 4 LysLeuThrGluGluGln1leSerGluPheLysGluAlaPheSerLeuPheAsp 23
 QY 490 GAGATGCTTCTGCTGATGTCGAGAGAGTGAATGTTGCTCCAGAGAAATCGGG----- 437
 DB 24 GlyAspGlyThrLeuThrLysGluLeuGlyThrValMetArgSerLeuGlyGlnAen 43
 QY 436 -----TCTGTCATTCACGCTTGAAGAAATGAGAGCCCTTACTGGGAAACCA 389
 DB 44 ProThrGluAlaGluLeuGlnAspMet1leAsnGlu1leAsp---ThrAspGlyAsnGly 62
 QY 388 TGTTTGAATCTGGATGAATCTTCTTCTTGAATTCATCTCGAACCCATCGGCACAT 329
 DB 63 Thr1leAspPheProGluPheLeuThrLeu-----Met1laArgLysLeuLys--- 78
 QY 328 GTGTGTCGACGAAGAGAGAGAGAGATGTCATTCACGCGCGCGTGAAGAAAGAC 269
 DB 79 -----AspThrAspPheThrGluGluLeu1le----- 87
 QY 268 AGTACCTTCGAGAGCTTTTAAAGTGTGACTTGAATGGAGATGGTGGGAGATGT 209
 DB 88 -----GluAlaPheArgValPheAspArgAspGlyTyrl1leSerAla 103
 QY 208 GAGAGCTTGAATACGCTGCGGAGAGACTGGGTATGTGAGTGAATAATGTGAAAAAGAC 149
 DB 104 AspGluLeuArgHisValMetThrAsnLeuGly-----GluLysLeuThrAsnGluGlu 121
 QY 148 TGCAGAGAGATGATTTGGTATTACGACACCAATTCAGAGCGCATGTTGTTCAAGAA 89
 DB 122 ValAspGluMet1leArgGluAlaAsp1leAspGlyAspGlyGln1leAsnGlyGluGlu 141
 QY 88 TTTCAAAAATCATGATTGA 71
 DB 142 PheValLysMetMet1le 147
 RESULT 7
 CALM_STYLE STANDARD; PRT; 148 AA.
 ID CALM_STYLE P27166;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Calmodulin.
 OS Stylonychia lemnae.
 OC Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
 OC Stichotrichida; Oxytrichidae; Stylonychia.
 OX NCBI_TaxID=5949;
 RN [1]
 RP SEQUENCE FROM N.A.

QY 310 GAGAGAGATTGTCATTCACGGCGGGTGAAGAAGACGACCTTGGAAGCT 251
 Db 84 SerGluGlu-----GluLeuArgGluAla 91
 QY 250 TTTAAAGCTTTGACTGTAATGGGATGGGGGATGTGAGAGAGCTTGAAATCGTG 191
 Db 92 PheIysValAlaPheAspIysAspGlyIleIleSerAlaIleGluLeuArgHisVal 111
 QY 190 CTGGGAAGACTGGGTATGTGAGAGTGAATAATAGTGAAGAAAGACTGCGAGACATGATTTGG 131
 Db 112 MetThrAsnLeuGly-----GluYsLeuThrAspGluIleValAspGluMetIleArg 129
 QY 130 TATTACGACACCATTCAGACGCGATGTTGATTTTCAGAAATTCAGAAATGATGTTA 71
 Db 130 GluAlaAspValAspIysAspGlyValIleAspTyrSerGluPheValIysMetIleLeu 149
 RESULT 11
 TPCS MOUSE
 ID TPCS MOUSE STANDARD; PRT; 159 AA.
 ~ 20801;
 01-FEB-1991 (Rel. 17, Created)
 01-FEB-1991 (Rel. 17, Last sequence update)
 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Tropoin C, skeletal muscle (STNC).
 GN TNMC2 OR TNCS.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BAH/c;
 RX MEDLINE=90368819; PubMed=2394755;
 RA Parmacek M.S., Bengur A.R., Vora A.J., Leiden J.M.;
 RT "The structure and regulation of expression of the murine fast
 RT skeletal tropoin C gene. Identification of a developmentally
 RT regulated, muscle-specific transcriptional enhancer";
 RL J. Biol. Chem. 265:15970-15976(1990).
 CC -1- FUNCTION: TROPONIN IS THE CENTRAL REGULATORY PROTEIN OF STRIATED
 CC MUSCLE CONTRACTION. TN CONSISTS OF THREE COMPONENTS: TN-I WHICH
 CC IS THE INHIBITOR OF ACTOMYOSIN ATPASE, TN-T WHICH CONTAIN THE
 CC BINDING SITE FOR TROPOMYOSIN AND TN-C, THE BINDING OF CALCIUM TO
 CC TN-C ABOLISHES THE INHIBITORY ACTION OF TN ON ACTIN FILAMENTS.
 CC -1- TISSUE SPECIFICITY: FAST SKELETAL MUSCLE.
 CC -1- MISCELLANEOUS: SKELETAL MUSCLE TROPONIN C BINDS FOUR CALCIUM IONS.
 CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL_M57590; AAA37642.1; -
 DR PIR; A38383; A38383.
 DR HSSP; P02586; 1A2X.
 DR SWISS-2DPAGE; P20801; MOUSE.
 DR MGD; MGI:98780; Tncs.
 DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; ehand; 4.
 DR ProDom; PD000012; EF-hand; 2.
 DR SMART; SM00054; EFh; 4.
 DR ProSITE; PS00018; EF_HAND; 4.
 KM Muscle protein: Calcium-binding. Repeat; Acetylation.
 FT INT MET 0
 FT MOD_RES 1 1
 FT CA_BIND 27 38 EF-HAND 1.
 FT CA_BIND 63 74 EF-HAND 2.
 FT CA_BIND 103 114 EF-HAND 3.
 FT CA_BIND 139 150 EF-HAND 4.
 SEQUENCE 159 AA; 17979 MW; 65B1A9EECC06023 CRC64;

Alignment Scores:
 Pred. No.: 6 21e-06 Length: 159
 Score: 154.50 Matches: 43
 Percent Similarity: 45.86% Conservative: 29
 Best Local Similarity: 27.39% Mismatches: 60
 Query Match: 14.56% Indels: 25
 Gaps: 4
 DB: 1
 US-10-021-323-13 (1-609) x TPCS_MOUSE (1-159)
 QY 526 GACTTGCACACCGGTATTCGAGAGAGCTCGACAAAGATGAGATGCTTGTAGCTGGAG 467
 Db 18 GluPheIysAlaAlaPheAspMetPheAspAlaAspGlyIleAspIleSerValIys 37
 QY 466 GAGCTGAATGCTGCTCCAGAGAAATCGGGTCTGTCCAAATTCAGCTTGAAGATTTGAG 407
 Db 38 GluLeuGlyThrValMetArgMetLeuGlyGlnThrPro---ThrIleGluIleuAsp 56
 QY 406 CCCTTAGTG-----GGAACCATGTTGAACTTGATGATGATCTTGT 365
 Db 57 AlaIleIleGluGluValAspIleAspIleSerGlyThrIleAspPheGluIleuAsp 76
 QY 364 TTCTTTTATGATTCATCTCGAACCCACTGCGACATGTTGTTGACGAAAGAGAGAGAG 305
 Db 77 ValMetMetValArgGluMetIleGluAspAlaIysGlyIleSerGluIleu----- 94
 QY 304 GAATGTGATTCATTCACGGCGCGGTGAAGAAGACAGTGCCTTGCAGAGCTTTTAA 245
 Db 95 -----LeuAlaGluCysPheArg 100
 QY 244 GTGTTGACTGATGATGGGATGGTGGGGGATGTGAGAGCTTGAATACGCTGTGGA 185
 Db 101 IlePheAspArgAsnAlaAspIlyIleAspAlaGluGluIleuAlaGluIlePheArg 120
 QY 164 AGACTGGGTATCTGAGAGTGAATAATGTGGAAAGCTCGACAGATGATTTGTTATAC 125
 Db 121 AlaSerGly-----GluHisValThrIleGluIleGluSerLeuMetIleAspIly 138
 QY 124 GACACCAATTCAGACGCGATGTTGATTTTCAGAAATTCAGAAATCATGATG 74
 Db 139 AspIysAsnAsnAspIlyArgIleAspPheAspIleuIleuMet 155
 RESULT 12
 CALM_EUGER
 ID CALM_EUGER STANDARD; PRT; 148 AA.
 AC P1118;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE Calmodulin.
 OS Euglena gracilis.
 OC Eukaryota; Euglenozoa; Euglenida; Euglenales; Euglena.
 OX NCBI_TaxID=3039;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=2;
 RX MEDLINE=92241300; PubMed=1572365;
 RA Toda H., Yazawa M., Yagi Y.;
 RT "Amino acid sequence of calmodulin from Euglena gracilis";
 RL Eur. J. Biochem. 205:653-660(1992).
 CC -1- FUNCTION: CALMODULIN MEDIATES THE CONTROL OF A LARGE NUMBER OF
 CC ENZYMES BY CA(++) AMONG THE ENZYMES TO BE STIMULATED BY THE
 CC CALMODULIN-CA(++) COMPLEX ARE A NUMBER OF PROTEIN KINASES AND
 CC PHOSPHATASES.
 CC -1- MISCELLANEOUS: THIS PROTEIN HAS FOUR FUNCTIONAL CALCIUM-BINDING
 CC SITES.
 CC -1- SIMILARITY: CONTRAINS 4 EF-HAND CALCIUM-BINDING DOMAINS.
 CC PIR; JK0008; MCEG.
 DR PIR; S21212; S21212.
 DR HSSP; P02593; 1CDM.
 DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; ehand; 4.

DR PRODOM; PD000012; EF-hand; 2.
 DR SMART; SM00054; EFh; 4.
 DR PROSITE; PS00018; EF_HAND; 4.
 KW Calcium-binding; Repeat; Acetylation; Methylation.
 FT MOD_RES 1 1 ACETYLATION.
 FT MOD_RES 115 115 METHYLATION.
 FT MOD_RES 148 148 METHYLATION.
 FT MOD_RES 20 31 EF-HAND 1.
 FT CA_BIND 56 67 EF-HAND 2.
 FT CA_BIND 93 104 EF-HAND 3.
 FT CA_BIND 129 140 EF-HAND 4.
 SQ SEQUENCE 148 AA; 16724 MW; 82A1E48108638455 CRC64;

Alignment Scores:
 Pred. No.: 6.8e-06 Length: 148
 Score: 154.00 Matches: 43
 Percent Similarity: 49.064 Conservative: 35
 Best Local Similarity: 27.044 Mismatches: 51
 Query Match: 14.518 Indels: 30
 Gaps: 6

us-10-021-323-13 (1-609) x CALM_EUGR (1-148)

QY 526 GACCTGCAACGGCGTATTGAGAAAGCTCGACAGAAATGAGATGCTTCTGCTGAG 467
 Db 11 GluphelysGlualphaserleupheaplysaspolyThrillethrlhrls 30
 QY 466 GAGCGAATGGTGTGTCAGAGAAATCGGCGTCTGTCATTCAGCTTGA--GAATG 410
 Db 31 GluleGlyThrValmetaraserleugly-----GlnsbmProthrlalaglulbe 48
 QY 409 GAGCCTTAGTG-----GGAAACCATGTTGAATCTGATGATTC 368
 Db 49 GlnasmetIleasndluvalaspllnasplyserGlyThrileasphProdiuphe 68
 QY 367 TTGTTCTTTATGATTCATCTCGAACCCACTGGCACTGATGATGATGATGATGATG 308
 Db 69 Leu-----ThrluwtSerArglyGlyMetHlsapThrsapThrgluGlu 84
 QY 307 GAGGAATGATGATTCATCAGCGCGCGGTGAAGAAGACAGTACCTGCGAAGCTTT 248
 Db 85 -----Ileysglualaspe 89
 QY 247 AAAGTGTGACTGAATGAGGATGGGCGGATGTGAGAGCTTGATATCGTCTG 188
 Db 90 ArgvalPheaspyaspIysaspIysaspIleSeraIalagluleuArgHlsValme 109
 187 GGAAGCTGGGTATGTGAGGTGAATAGTGGAAAAGACTCGAGAGCATGATTGGTAT 128
 110 ThrAsnleugly-----GluYsLeuThrAspGluGluValaspllnuettIleArgGlu 127
 QY 127 TACGACACCAATCAGACGCGATGTTGATTCAAGAAATCAAAACATGATGTA 71
 Db 128 AlaasvalaspolysaspIysaspIysaspIleasentryGluGluValasyspMetMet 146

RESULT 13
 TPCS CHICK STANDARD; PRT; 162 AA.
 AC P02588;
 DT 21-JUN-1986 (Rel. 01, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Troponin C, skeletal muscle.
 GN TNNG2.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RX MEDLINE=86115393; PubMed=2963002;
 RA Reinach F.C., Karlsson R.;

RT "Cloning, expression, and site-directed mutagenesis of chicken
 RT skeletal muscle troponin C.";
 RL J. Biol. Chem. 263:2371-2376(1988).
 RN (12)
 RP SEQUENCE.
 RX MEDLINE=93340719; PubMed=1908459;
 RA Golosinski K., Pearlstone J.R., Borgford T., Oikawa K., Kay C.M.,
 RA Carpenter M.R., Smillie L.B.;
 RT "Determination of and corrections to sequences of turkey and chicken
 RT troponin-C. Effects of Thr-130 to Ile mutation on Ca2+ affinity.";
 RL J. Biol. Chem. 266:15797-15809(1991).
 RN (3)
 RP PRELIMINARY SEQUENCE.
 RX MEDLINE=77048905; PubMed=992069;
 RA Wilkinson J.M.;
 RT "The amino acid sequence of troponin C from chicken skeletal muscle.";
 RL FEBS Lett. 70:254-256(1976).
 RN (4)
 RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
 RX MEDLINE=85115391; PubMed=3969570;
 RA Sundaralingam M., Bergstrom R., Straaburg G., Rao S.T.,
 RA Raychowdhury P., Greaser M.L., Wang B.C.;
 RT "Molecular structure of troponin C from chicken skeletal muscle at 3-A.
 RT resolution.";
 RL Science 227:945-948(1985).
 RN (5)
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RX MEDLINE=88115277; PubMed=3338985;
 RA Sanyshur K.A., Rao S.T., Pyzalska D., Drendel W., Greaser M.L.,
 RA Sundaralingam M.;
 RT "Refined structure of chicken skeletal muscle troponin C in the two-
 RT calcium state at 2-A resolution.";
 RL J. Biol. Chem. 263:1628-1647(1988).
 RN (6)
 RP X-RAY CRYSTALLOGRAPHY (1.78 ANGSTROMS).
 RA Sanyshur K.A., Pyzalska D., Rao S.T., Greaser M.L., Sundaralingam M.;
 RT "Structure of chicken skeletal muscle troponin C at 1.78-A
 RT resolution.";
 RL Acta Crystallogr. D 50:40-49(1994).
 RN (7)
 RP X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS).
 RX MEDLINE=98035054; PubMed=9367759;
 RA Strynadka N.C., Cherney M., Stielecki A.R., Li M.X., Smillie L.B.,
 RA James W.N.G.;
 RT "Structural details of a calcium-induced molecular switch: X-ray
 RT crystallographic analysis of the calcium-saturated N-terminal domain
 RT of troponin C at 1.75-A resolution.";
 RL T. Mol. Biol. 273:238-255(1997).
 RN (8)
 RP STRUCTURE BY NMR OF 93-126.
 RX MEDLINE=93003111; PubMed=1390738;
 RA Shaw G.S., Hodges R.S., Sykes B.D.;
 RT "Determination of the solution structure of a synthetic two-site
 RT calcium-binding homodimeric protein domain by NMR spectroscopy.";
 RL Biochemistry 31:9572-9580(1992).
 RN (9)
 RP STRUCTURE BY NMR.
 RX MEDLINE=96101426; PubMed=8519752;
 RA Slupeky C.M., Sykes B.D.;
 RT "NMR solution structure of calcium-saturated skeletal muscle troponin
 RT C.";
 RL Biochemistry 34:15953-15964(1995).
 RN (10)
 RP STRUCTURE BY NMR.
 RX MEDLINE=99249778; PubMed=10231519;
 RA Tsuda S., Miura A., Gagne S.M., Spyropoulos L., Sykes B.D.;
 RT "Low-temperature-induced structural changes in the Apo regulatory
 RT domain of skeletal muscle troponin C.";
 RL Biochemistry 38:5693-5700(1999).
 CC -1- FUNCTION: TROPONIN IS THE CENTRAL REGULATORY PROTEIN OF STRIATED
 CC MUSCLE CONTRACTION. TN CONSISTS OF THREE COMPONENTS: TN-I WHICH
 CC IS THE INHIBITOR OF ACTOMYOSIN ATPASE, TN-T WHICH CONTAIN THE
 CC BINDING SITE FOR TROPOMYOSIN AND TN-C. THE BINDING OF CALCIUM TO

CC TN-C ABOLISHES THE INHIBITORY ACTION OF TN ON ACTIN FILAMENTS.
CC -1- MISCELLANEOUS: SKELETAL MUSCLE TROPONIN C BINDS FOUR CALCIUM IONS.
CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-stb.ch/announce/>
CC or send an email to license@isb-stb.ch).
CC -----
CC EMBL, M19027; AAA49097.1; ALT_SEQ.
DR PIR, A03015; TPCHS.
DR PDB, 4TNC; 19-APR-89.
DR PDB, 1TOP; 31-JAN-94.
DR PDB, 1NCP; 07-DEC-96.
DR PDB, 1NCK; 07-DEC-96.
DR PDB, 1NCG; 11-JAN-97.
DR PDB, 1NCH; 07-DEC-96.
DR PDB, 1POM; 08-NOV-96.
DR PDB, 1SMG; 12-AUG-97.
DR PDB, 1TNP; 15-OCT-95.
DR PDB, 1TNO; 15-OCT-95.
DR PDB, 1TNW; 15-OCT-95.
DR PDB, 1TNX; 15-OCT-95.
DR PDB, 1AVS; 24-DEC-97.
DR PDB, 1ZAG; 11-NOV-98.
DR PDB, 1SKT; 13-JAN-99.
DR PDB, 1BLQ; 13-JAN-99.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 4.
DR ProDom; PD000012; EF-hand; 2.
DR SMART; SM00054; EFh; 4.
DR PROSITE; PS00018; EF-HAND; 4.
KM Muscle protein; Calcium-binding; Repeat; 3D-structure.
FT INIT MET 0
FT MOD RES 1
FT CA_BIND 30 41 EF-HAND 1.
FT CA_BIND 66 77 EF-HAND 2.
FT CA_BIND 106 117 EF-HAND 3.
FT CA_BIND 142 153 EF-HAND 4.
FT MUTAGEN 130 130 T->I. DECREASES CALCIUM AFFINITY.
FT HELIX 3 13
FT HELIX 16 27
FT TURN 28 30
FT STRAND 36 38
FT HELIX 39 48
FT TURN 49 50
FT HELIX 55 65
FT TURN 67 68
FT STRAND 72 74
FT HELIX 75 105
FT TURN 107 108
FT STRAND 113 113
FT HELIX 115 123
FT TURN 124 126
FT HELIX 131 141
FT TURN 143 144
FT STRAND 149 149
FT HELIX 151 159
SQ SEQUENCE 162 AA; 18244 MW; 9E2EDF0B5A3EDA4C CRC64;

Alignment Scores:
Pred. No.: 7.61e-06 Length: 162
Score: 153.50 Matches: 44
Percent Similarity: 44.59% Conservative: 26
Best Local Similarity: 28.03% Mismatches: 62
Query Match: 14.47% Indels: 25
DB: 1 Gaps: 4

US-10-021-323-13 (1-609) x TPCS_CHICK (1-162)
QY 526 GACCTGCAACGCGATTGAGAGCTCGACAAGATGAGATGCTGCTAGTCTGGAG 467

Db 21 GluphetyalaAlaPheaspMetPheaspAlaAspGlyGlyAspIleSerThrLys 40
QY 466 GACCTGGAATGGTTCCTCCAGAAATCCGGCTCTGTCATTCACGCTTGAAGATTGGAG 407
Db 41 GlueuglythrValMetLArgMetLeuGlyGln--AenProThrLysGluGluLeuAsp 59
QY 406 CCCTTAGTG-----GGAAACCATGTTGACCTTGATGATGATTTG 365
Db 60 AlaIleIleGluGluValAlaAspLysAspIleSerGlyThrIleAspPhcIuGluPheLeu 79
QY 364 TTTCTTTATGATTCATCTGCAACCCACTGSCACATGTGTGTGACGAAGAGAGAGAG 305
Db 80 ValMetMetValArgGluMetLysGluAspLysGlyLysSerGluGluGln----- 97
QY 304 GAATTGTCATTCACGGGGGGGTGAAGAAGAAGACATGACCTTGCGAAGCTTTTAA 245
Db 98 -----LeuAlaAsnCysPheArg 103
QY 244 GTGTTTGACTTGAATGGGGATGGTGGGGGATGTGAGAGACTTGAATACGCTGGGA 185
Db 104 IlePheAspLysAsnAlaAspGlyPheIleAspIleGluGluLeuGlyGluLeuArg 123
QY 184 AGACTGGGTATGTAGTGAAGTGAATAAGTGAAGAAGACTGCAGAGCATGATTGTATAC 125
Db 124 AlaThrGly-----GluHleValThrGluGluAspIleGluAspLeuMetLysAspSer 141
QY 124 GACACCAATTCAGACGGCATGTTGATTTTCAAGATTCAAAACATGATG 74
Db 142 AspLysAsnAsnAspGlyArgIleAspPheAspGluPheLeuLysMet 158

RESULT 14
TPCS_HUMAN
ID TPCS_HUMAN STANDARD; PRT; 159 AA.
AC P02585;
DT 21-JUL-1986 (Rel. 01, Created)
DR 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Troponin C, skeletal muscle.
GN TNNC2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88332973; PubMed=3166492;
RA Gahlmann R., Wade R., Gunning R., Kedes L.;
RT "Differential expression of slow and fast skeletal muscle troponin C.
RT Slow skeletal muscle troponin C is expressed in human fibroblasts."
RL J. Mol. Biol. 201:379-391(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90324243; PubMed=2373703;
RA Gahlmann R., Kedes L.;
RT "Cloning, structural analysis, and expression of the human fast
RT twitch skeletal muscle troponin C gene."
RL J. Biol. Chem. 265:12520-12528(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Mu Q.L.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Copley V.E., Collier R.E., Connor R.E., Corby N.R.,

RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Griffith D.V., Griffiths C., Griffiths M.N.D., Gilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.W., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Leishvala M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McLeay K., Murray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mulliken J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Philimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H.,
RA Rice C.M., Ross M.T., Scott C.E., Senra H.K., Showkhen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.,
RT "The DNA sequence and comparative analysis of human chromosome 20.",
Nature 414:865-871(2001).
[5]
SEQUENCE.
RX MEDLINE=77031728, PubMed=978749;
RA Romero-Herrera A.E., Castillo O., Lehmann H.,
RT "Human skeletal muscle proteins. The primary structure of troponin
C.",
J. Mol. Evol. 8:251-270(1976).
CC -1- FUNCTION: TROPONIN IS THE CENTRAL REGULATORY PROTEIN OF STRIATED
CC MUSCLE CONTRACTION. TN CONSISTS OF THREE COMPONENTS: TN-I WHICH
CC IS THE INHIBITOR OF ACTOMYOSIN ATPASE, TN-T WHICH CONTAIN THE
CC BINDING SITE FOR TROPOMYOSIN AND TN-C. THE BINDING OF CALCIUM TO
CC TN-C ABOLISHES THE INHIBITORY ACTION OF TN ON ACTIN FILAMENTS.
CC -1- MISCELLANEOUS: SKELETAL MUSCLE TROPONIN C BINDS FOUR CALCIUM IONS.
CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X07898; CAA30737.1; -
DR EMBL: M33772; AAA61197.1; -
DR EMBL: M33771; AAA61197.1; JOINED.
DR EMBL: M22307; AAA91854.1; -
DR EMBL: AL050348; CAB66117.1; -
DR PIR: A03012; TPBCS.
DR PIR: B29990; B29990.
DR PIR: A36574; A36574.
DR HSSP: P02586; IAXX.
DR GeneW: HGNC:11944; TNNC2.
DR MIM: 191039; -
DR InterPro: IPR002048; EF-hand.
DR Pfam: PF00036; efhand; 4.
DR ProDom: PD00012; EF-hand; 2.
DR SMART: SM00054; EFh; 4.
DR PROSITE: PS00018; EF_HAND; 4.
DR Muscle protein; Calcium-binding; Repeat; Acetylation.
DR KW INIT MET 0
FT MOD RES 1 ACETYLATION.
FT CA_BIND 27 38 EF-HAND 1.
FT CA_BIND 63 74 EF-HAND 2.
FT CA_BIND 103 114 EF-HAND 3.
FT CA_BIND 139 150 EF-HAND 4.
FT CONFLICT 1 TD -> DT (IN REF. 5).
FT CONFLICT 113 2 E -> G (IN REF. 3).
SQ SEQUENCE 159 AA, 17991 MW, 1506B29ECGCC06036 CRC64;

Alignment Scores:

Pred. No.: 9,29e-06 Length: 159
Score: 152.50 Matches: 43

Percent Similarity: 45.86% Conservative: 29
Best Local Similarity: 27.39% Mismatches: 60
Query Match: 14.37% Indels: 25
DB: 1 Gaps: 4
US-10-021-323-13 (1-609) x TPCC_HUMAN (1-159)
QY 526 GACCTTGACACCGCGATTCGAGAACCTGACAGAAATGAGATGCTTCGTTAGTCTGAG 467
DB 18 GluHeuValAlaIlePheAspMetPheAspIleAspGlyGlyIleAspIleSerValIys 37
QY 466 GAGCTGAATGCTGCTCCACAGAAATGGGCTCTCCATTCCATTACCTTGAAGATTGGAG 407
DB 38 GluHeuGlyThrValMetArgMetLeuGlyInThrPro--ThrIysGluGluLeuAsp 56
QY 406 CCCTTAGG-----GGAACACCATGTTGAATGATGATGATTTCTG 365
DB 57 AlaIleIleGluGluValAspGluAspIlyserGlyThrIleAspPheGluGluPheLeu 76
QY 364 TTCTTTATGATGATCCATCTCGAACCCACCTGCACATGCTGTGACGAGAGAGAGAG 305
DB 77 ValMetMetValArgGluMetIleGluAspIleAlaIysGlyLysSerGluGluGlu----- 94
QY 304 GAATGCTGATTCACGCGCGCGGTGAAGAGACACTGACCTTGGAGAGCTTTTAA 245
DB 95 -----LeuAlaGluCysPheArg 100
QY 244 GTGTTTGACTGAATGGGAGATGGGTGGGGGATGTGAGAGCTTGAATACGTGCTGGGA 185
DB 101 IlePheAspArgAlaIleAspGlyIlyrIleAspProGluGluLeuAlaGluIlePheArg 120
QY 184 AGACTGGGTATGATGAGTGAATAATGAGAAAGACTGTCAGAGACATGATTTGGATTAC 125
DB 121 AlaSerGly-----GluHisValThrAspGluGluIleGluSerLeuMetIleAspIly 138
QY 124 GACACCAATGACAGCGCATGCTGATTTTCAAGATTCAAAAATGATG 74
DB 139 AspIysAsnAsnAspGlyArgIleAspPheAspGluPheLeuIyMetMet 155
RESULT 15
TPCC_RANES STANDARD; PRT; 162 AA.
AC P02589;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Troponin C, skeletal muscle.
OS Rana esculenta (Edible frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.
OX NCBI_TaxID=8401;
RN [1]
RP SEQUENCE.
RX MEDLINE=79065043; PubMed=309817;
RA van Eerd J.-P., Capony J.-P., Ferraz C., Pecheze J.-F.,
RT "The amino-acid sequence of troponin C from frog skeletal muscle.",
RL Eur. J. Biochem. 91:231-242(1978).
CC -1- FUNCTION: TROPONIN IS THE CENTRAL REGULATORY PROTEIN OF STRIATED
CC MUSCLE CONTRACTION. TN CONSISTS OF THREE COMPONENTS: TN-I WHICH
CC IS THE INHIBITOR OF ACTOMYOSIN ATPASE, TN-T WHICH CONTAIN THE
CC BINDING SITE FOR TROPOMYOSIN AND TN-C. THE BINDING OF CALCIUM TO
CC TN-C ABOLISHES THE INHIBITORY ACTION OF TN ON ACTIN FILAMENTS.
CC -1- MISCELLANEOUS: SKELETAL MUSCLE TROPONIN C BINDS FOUR CALCIUM IONS.
CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
DR PIR: A03016; TPBCS.
DR HSSP: P10246; ITRF.
DR InterPro: IPR002048; EF-hand.
DR Pfam: PF00036; efhand; 4.
DR ProDom: PD00012; EF-hand; 2.
DR SMART: SM00054; EFh; 4.
DR PROSITE: PS00018; EF_HAND; 4.
DR Muscle protein; Calcium-binding; Repeat; Acetylation.
DR KW MOD_RES 1 ACETYLATION.
FT

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_nzp model

Run on: June 24, 2003, 13:57:57 ; Search time 48 Seconds

(without alignments)
5228.446 Million cell updates/sec

Title: US-10-021-323-13

Perfect score: 1100

Sequence: 1 ggtatgcaatcaactctt.....tgaagcctacaataaag 609

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

arched: 671580 seqs, 20604715 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first: 45 summaries

Command line parameters:

-MODEL=frame+np.model -DEV=xlp
-Q=/cgn2.1/USPRO.spool/US10021333/runat_24062003.102234.20482/app.query.fasta_1.775
-DB=SPTRMBL 21 -OPMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -STRAP=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USR=US10021333 @CGN 1 1 138 @runat_24062003.102234.20482 -NCPU=6 -ICPU=3
-NO MMAP -LARGEDBRY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPTRMBL 21:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phage:
10: sp_plant:
11: sp_prodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_virus:
16: sp_bacterioid:
17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being pinned, and is derived by analysis of the total score distribution.

Result No. Score Match Length DB ID Description
C 1 312 29.4 142 10 Q9LPK5 Q9LPK5 arabidopsis

C	2	275.5	26.0	191	10	Q9SVG9	Q9svg9 arabidopsis
C	3	265	25.0	150	10	Q9SCA1	Q9scA1 lotus japon
C	4	245.5	23.1	150	10	Q9LME7	Q9lme7 arabidopsis
C	5	241.5	22.8	215	10	Q92845	Q92845 arabidopsis
C	6	237.5	22.4	154	10	Q92R02	Q92R02 arabidopsis
C	7	234	22.1	152	10	Q9SU00	Q9su00 arabidopsis
C	8	227	21.4	153	10	Q9SRR7	Q9srr7 arabidopsis
C	9	218.5	20.6	181	10	Q9F109	Q9f109 arabidopsis
C	10	218	20.5	205	10	Q9FR00	Q9fr00 arabidopsis
C	11	216.5	20.4	195	10	Q9LX27	Q9lx27 arabidopsis
C	12	208.5	19.7	185	10	Q9R2B5	Q9r2b5 oryza sativ
C	13	206.5	19.5	172	10	Q93Y88	Q93y88 sesbania ro
C	14	204	19.2	145	5	Q09980	Q09980 caenorhabdi
C	15	201.5	19.0	188	10	Q9ARP2	Q9arp2 penitsemu
C	16	200	18.9	167	10	Q9AR93	Q9ar93 medicago sa
C	17	199	18.8	163	10	Q9C9U8	Q9c9u8 arabidopsis
C	18	194	18.3	190	10	Q9AXG2	Q9axg2 nicotiana t
C	19	192	18.1	157	10	Q9C8Y1	Q9c8y1 arabidopsis
C	20	191	18.0	186	10	Q9FYK2	Q9fyk2 arabidopsis
C	21	184.5	17.4	144	10	Q9SJN6	Q9sjn6 arabidopsis
C	22	182	17.2	166	10	Q94SR0	Q94sr0 oryza sativ
C	23	178	16.8	170	10	Q9LE22	Q9le22 arabidopsis
C	24	175.5	16.5	146	10	Q8RYJ8	Q8ryj8 oryza sativ
C	25	174	16.4	134	5	P90620	P90620 trichomonas
C	26	173.5	16.4	187	10	Q9Z0E6	Q9z0e6 arabidopsis
C	27	171	16.1	156	5	Q18136	Q18136 caenorhabdi
C	28	170	16.0	321	10	Q943Q4	Q943q4 oryza sativ
C	29	169	15.9	146	10	Q9LFE5	Q9lfe5 arabidopsis
C	30	167	15.7	173	10	Q93WY1	Q93wy1 musa acumin
C	31	165.5	15.6	151	10	Q8RYJ9	Q8ryj9 oryza sativ
C	32	164	15.5	164	10	Q39708	Q39708 dunaliella
C	33	163.5	15.4	185	10	Q9FTH9	Q9fth9 arabidopsis
C	34	163	15.4	159	10	Q9SKR7	Q9skr7 arabidopsis
C	35	161	15.2	204	10	Q8R2B7	Q8r2b7 oryza sativ
C	36	160.5	15.1	149	5	Q95V99	Q95v99 leishmania
C	37	160.5	15.1	150	10	Q40982	Q40982 pismu sativ
C	38	159	15.0	151	10	Q8RYK0	Q8ryk0 arabidopsis
C	39	158	14.9	151	5	Q96792	Q96792 brachiosteo
C	40	158	14.9	575	10	Q9ZSA3	Q9zsa3 arabidopsis
C	41	156	14.7	149	11	Q9D6P8	Q9d6p8 mus musculu
C	42	155	14.6	150	10	Q39890	Q39890 glycine max
C	43	154	14.5	138	10	Q15931	Q15931 symbiodinu
C	44	154	14.5	149	11	Q99K52	Q99k52 mus musculu
C	45	153	14.4	149	5	Q96102	Q96102 physarum po

ALIGNMENTS

RESULT 1
Q9LPK5 PRELIMINARY; PRT; 142 AA.
AC Q9LPK5;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE F2408.15 protein (Act121550/F2408_7).
GN F2408.15.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eudicots II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Federapfel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Alarafi H., Nguyen M., Lam B., Southwick A., Miranda M., Brooks S.,
RA Buehler E., Chao Q., Chin C., Chou J., Choi E., Gonzalez A.,
RA Huang B., Johnson-Hopson C., Khan S., Kim C., Koo T., Lee J.M.,
RA Lenz C., Liu A., Liu S., Mukharay N., Pham P., Sakano H., Shim P.,
RA Tortum M., Vaysberg M., Yu G., Ecker J., Theologis A., Davis R.W.,
RA Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

RA Shlun P., Chen H., Cheuk R., Kim C.J., Banh J., Bowser L.,
 RA Carinini P., Chung M.K., Goldsmith A.D., Hayashizaki Y., Ishida J.,
 RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M.,
 RA Lin J., Liu S.X., Mitranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M., Seki M.,
 RA Southwick A., Tortum M., Yamada K., Yu G., Shinozaki K., Davis R.W.,
 RA Theologis A., Ecker J.R.,
 RT "Arabidopsis cDNA clones",
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC015447; AAF87905.1; -
 DR EMBL; AF161594; AKK32762.1; -
 DR HSSP; P02593; ICDM.
 DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; ehand; 3.
 DR Prodom; PD000012; EF-hand; 1.
 DR SMART; SM00054; EFh; 3.
 DR PROSITE; PS00018; EF HAND; UNKNOWN 2.
 SQ SEQUENCE 142 AA; 1638 MW; 339C056754E23AB9 CRC64;

Alignment Scores:
 d. No.: 1,486-23 Length: 142
 Score: 312.00 Matches: 68
 Percent Similarity: 63.33% Conservative: 27
 Best Local Similarity: 45.33% Mismatches: 45
 Query Match: 29.41% Indels: 10
 DB: 10 Gaps: 3

US-10-021-323-13 (1-609) x Q9SVG9 (1-142)
 QY 514 GTATTCGAGAGAGCTGACAGAGATGAGATGCTGCTGATGAGAGAGCTGAATTGG 455
 Db 1 MetPheIyStHLeuAspIysAsnGInAspGlyLeuValThrLeuAspGluLeuLeuTyr 20
 QY 454 TTGCTCCAGAGAGATGGGGCTGCTGCAATTCAGCTTGAAGATTGAGCCCTTAGTGGGA 395
 Db 21 ILeuAspIysLeuGlyTyrPalagIuHstHrProAspGluLeuGluLeuValGly 40
 QY 394 AAACCATGTTGAAGCTTGATGATGATTTCTG--TTCTTTATGAATCCATCTGCAACCA 338
 Db 41 LysGInSerLeuAspLeuAspGluPheLeuArgPheTyrTyrAspAlaVal----- 57
 QY 337 CTGGACATGCTGTGACGACAGAGAGAGAGAAATTGCTCATTCACGGCGCGGTAA 278
 Db 58 LeuAspSerIysGlySerIysIysAsnIleAspValAla----- 71
 QY 277 GAAGAAGACAGTCACTTGCTGCAAGCTTTAAAGTGTGATGATGAAATGGGATGGTGG 218
 Db 72 AspaAspGluAlaIleAlaIleArgAlaPheAsnValPheAspValAsnGlyAspGlyTyr 91
 QY 217 GGGGATGTGAGAGCTTGATGATGATGCTGCGAGAGAGCTGGTATGTGAGTGAATAAGT 158
 Db 92 IleserAlaGluGluLeuArgAspValIleuGluArgLeuGlyPheGluGluGluAlaIys 111
 QY 157 GGAAGAGACTGACAGAGATGATTTGGTATTCAGACCAATTCAGACGCGCATGGTTGAT 98
 Db 112 AlaTPrAspCyGlyArgMetIleArgValHisAspIysAsnLeuAspGlyPheValAsp 131
 QY 97 TTTCAGAAATTCAAAAACATGATGTTACAT 68
 Db 132 PheGluGluPheIysAsnMetIleLeuHis 141
 RESULT 2
 Q9SVG9 PRELIMINARY; PRT; 191 AA.
 AC Q9SVG9;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Calcium-binding protein-like.
 GN P21C20.130 OR AT4G20780.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eutroside II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Beynon M., Polt T., Weisenegger T., Bancroft I., Mewes H.W.,
 RA Mayer K.F.X., Lemcke K., Schueller C.,
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN (2)
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN (3)
 RP SEQUENCE FROM N.A.
 RA Polt T., Weisenegger T., Mewes H.W., Lemcke K., Mayer K.F.X.,
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN (4)
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL080254; CAB45844.1; -
 DR EMBL; AL161553; CAB79078.1; -
 DR HSSP; P02593; ICDM.
 DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; ehand; 3.
 DR Prodom; PD000012; EF-hand; 2.
 DR SMART; SM00054; EFh; 3.
 DR PROSITE; PS00018; EF HAND; UNKNOWN 3.
 SQ SEQUENCE 191 AA; 21147 MW; 3B4FDEBCD4BC037 CRC64;

Alignment Scores:
 d. No.: 9,356-20 Length: 191
 Score: 275.50 Matches: 69
 Percent Similarity: 56.29% Conservative: 25
 Best Local Similarity: 41.32% Mismatches: 58
 Query Match: 25.97% Indels: 15
 DB: 10 Gaps: 5

US-10-021-323-13 (1-609) x Q9SVG9 (1-191)
 QY 544 TCCCCC-----CTTAGAAGACCGACTGACAGCGATTGAGAGAGCTGACAGAGATGA 488
 Db 22 SerProSerLeuAsnAlaLeuArgLeuGlnArgIlePheAspLeuPheAspIysAsnGly 41
 QY 487 GATGCTTCGTTAGTCTGAGAGAGCTGATGATGCTGCTCAAGATCGGCTGTCCTCA 428
 Db 42 AspGlyPheIleThrValGluGluLeuSerGlnAlaLeuThrArgLeuGly---LeuAsn 60
 QY 427 TTCACCTTGAAGAATTGGAGCCCTTAGTG-----GAAAAACCATGT 386
 Db 61 AlaAspLeuSerAspLeuIysSerThrValGluSerTyrIleGlnProGlyAsnThrGly 80
 QY 385 TTGAACCTTGAGATGATTTGTTGCTTTATGATGATTCATGCAACCACTG----- 335
 Db 81 LeuAsnPheAspPheSerSerLeuHisIysThrIleuAspAspSerPhePheGlyGly 100
 QY 334 GCACATGCTGTGAGAGAGAGAGAGAGAGAGAAATTGTCATTCACGGCGCGGTGAGAA 275
 Db 101 AlaCyGlyGlyGlyGluAsnGluAspAspPro-----SerSerAlaIleGlu 116
 QY 274 GAAGACAGTACCTTGCAAGGCTTTAAAGTGTGATCTGATGATGGGATGGGATGGGG 215
 Db 117 AenGluSerAspLeuAlaGluIleValPheIysValPheAspGluAsnGlyAspGlyPheIle 136
 QY 214 GGAATGTGAGAGAGCTGATGATGATGCTGGAGAGAGCTGGATGTAGTGAGTGAATAAGTGA 155
 Db 137 SerAlaArgGluLeuGlnThrValIleuIysLeuGlyLeuProGluGlyGlyGluMet 156
 QY 154 AAAGACTGACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATTT 95
 Db 157 GluArgValGluIysMetIleValSerValAspArgAsnGlnAspGlyArgValAspPhe 176
 QY 94 CAAGAATTCAAAAACATGATG 74

Db 177 PheGluPheLysAsnMetMet 183

RESULT 3

Q9SCA1 PRELIMINARY; PRT; 150 AA.

AC Q9SCA1: 09SCA1: PRELIMINARY; PRT; 150 AA.

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Calcium-binding protein (Fragment).

GN CBP1.

OS Lotus japonicus.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC Eurosid II; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.

OX NCBI_TaxID=34305;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. GIFU B-129-S9; TISSUE=ROOT;

ME MEDLINE=20289072; PubMed=10830260;

WE Webb K.J., Skot L., Nicholson M.N., Jorgensen B., Mizen S.; "Mesorhizobium loti increases root-specific expression of a calcium-binding protein homologue identified by promoter tagging in Lotus japonicus";

RT [2]. Plant Microbe Interact. 13:606-616(2000).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. GIFU B-129-S9; TISSUE=ROOT;

RU Skot L.;

DR EMBL: A251808; CAB63264.2; -.

DR HSSP: P02593; ICDW.

DR InterPro: IPR002048; EF-hand.

DR Pfam: PF00036; ehand; 4.

DR ProDom: PD000012; EF-hand; 2.

DR SMART: SM00054; Eph; 4.

DR PROSITE: PS00018; EF_HAND; UNKNOWN_4.

FT NON_TER 1

SQ SEQUENCE 150 AA; 17126 MW; 1286970A7FE73D71 CRC64;

Alignment Scores:

Pred. No.:	1 08e-18	Length:	150
Score:	265.00	Matches:	62
Percent Similarity:	55.90%	Conservative:	28
Best Local Similarity:	38.51%	Mismatches:	45
Query Match:	24.98%	Indels:	26
DB:	10	Gaps:	4

10-021-323-13 (1-609) x Q9SCA1 (1-150)

QY 538 CTTAGTAAGACCGACTTGCACCGCTATTGCGAAGAGCTCGACAAGATGAGATGCGCTTC 479

Db 1 MetAspProThrIleuLeuLysArgValPheGlnMetPheAspArgAsnGlyAspGlyArg 20

QY 478 GTTAGTCTGAGAGAGCTGAATGCTGCTCCAGAGAATCGGCTGTCTCAATTCAGCCTT 419

Db 21 IleThrIleLysGlnLeuAsnAspSerLeuGlnAsnLeuLys---IlePheIleProAsp 39

QY 418 GAAGATTGAGCGCTTAGTG-----GGAAACATGTTGAAGCTG 377

Db 40 LysGlnLeuThrGlnMetIleGluArgIleAspValAsnGlyAspGlyCysValAspIle 59

QY 376 GATGAATCTGTTGCTTTTATGATCATCTCGAACCCACTGGCAGCATGTGCTGACGAA 317

Db 60 AspGlnPheGlyGlnLeuLysArgIleMet-----AspGlu 72

QY 316 GAGGAGAGAGAGAAATGTCATTACGCGCGGGGTGAAGAGAAGACAGTACCTTGCG 257

Db 73 ArgAspGlnGluLys-----AspMetArg 80

QY 256 AAGCGTTTAAATGTTTGACTTGAATGGGATGGGTTGGGGGATGTGAGAGCTTGA 197

Db 81 GluAlaPheAsnValPheAspGlnAsnGlyAspGlyPheIleThrValGlnGluLeuArg 100

QY 196 TACGCTGGAAGACTGGTATGAGGTGAATAAGTGAAGAGCTGAGAGCATG 137

Db 101 ThrValLeuAlaSerIleuLysGlnGlyArgThrValGlnLysPylsLysMet 120

QY 136 ATTGGTATTACAGACACCAATTCAGCGGATGTTGATTTCAAGATTCAAAACATG 77

Db 121 IleMetLysValAspValAspGlyAspGlyMetValAspTyrLysGlnLysMet 140

QY 76 ATG 74

Db 141 Met 141

RESULT 4

Q9LNE7 PRELIMINARY; PRT; 150 AA.

AC Q9LNE7: 09LNE7: PRELIMINARY; PRT; 150 AA.

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE T2IE18.4 protein (putative calcium-binding protein).

GN T2IE18.4.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC Eurosid II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RU Sakano H., Vayenberg M., Lee J., Lenz C., Liu S.X., Pham P., Toriumi M., Yu G., Chin C., Chlou J., Choi E., Chung M., Gonzalez A., Howing B., Liu A., Alfati H., Brooks S., Buehler E., Chao Q., Conn L., Conway A.B., Hansen N.F., Johnson-Hopson C., Khan S., Kim C., Lam B., Miranda M., Nguyen M., Palm C.J., Shin P., Southwick A., Davis R.W., Becker J.R., Federjpiel N.A., Theologis A.;

RT "The sequence of BAC T2IE18 from Arabidopsis thaliana chromosome 1.";

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RU Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Egu P., Lee J.M., Toriumi M., Yu G., Brooks S., Chao Q., Chen H., Karlin-Neumann G., Kim C., Lam B., Miranda M., Nguyen M., Palm C.J., Shin P., Southwick A., Davis R.W., Becker J.R., Theologis A.;

RT "Full Length cDNA of gene T2IE18.4 (GI:8810461).";

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AC024174; AAF80122.1; -.

DR EMBL: AF332466; AAG48829.1; -.

DR HSSP: P02593; IFW4.

DR InterPro: IPR002048; EF-hand.

DR Pfam: PF00036; ehand; 4.

DR ProDom: PD000012; EF-hand; 2.

DR SMART: SM00054; Eph; 4.

DR PROSITE: PS00018; EF_HAND; UNKNOWN_4.

SQ SEQUENCE 150 AA; 16956 MW; 6AC33ADB5451D267 CRC64;

Alignment Scores:

Pred. No.:	1.11e-16	Length:	150
Score:	245.50	Matches:	58
Percent Similarity:	55.90%	Conservative:	32
Best Local Similarity:	36.02%	Mismatches:	46
Query Match:	23.14%	Indels:	25
DB:	10	Gaps:	4

US-10-021-323-13 (1-609) x Q9LNE7 (1-150)

QY 538 CTTAGTAAGACCGACTTGCACCGCTATTGCGAAGAGCTCGACAAGATGAGATGCGCTTC 479

Db 1 MetAspProThrIleuLeuLysArgValPheGlnMetPheAspArgAsnGlyAspGlyThr 20

QY 478 GTTAGTCTGAGAGAGCTGAATGCTGCTCCAGAGAATCGGCTGTCTCAATTCAGCCTT 419

Db 21 IleThrGlyLysGlnLeuLysArgIleMetPheAspArgSerLeuLys---IleTyrIleProAsp 39


```

OY 418 AAGAAATGAGACCCCTTACTG-----GGAAACCACTTTGAACCTTG 3177
Db 40 LysGluLeuThrGlmMetIleGluLysIleAspValAsnGlyCysValAspIle 59
OY 376 GATGAATCTCTGTTCTTTATGAAATTCATCTCGAACCCACTGGCACATGGTGTGACGAA 3178
Db 60 AspGluPheGlyGluLeuTyrTrpThrIleMet-----AspGlu 72
OY 316 GAGGAGGAGAGAGAGATGGTGCATTCACGGCGGGGTGAAGAGAAAGACACTGACCTTGGC 2573
Db 73 GluAspGluGluGlu-----GluAspMetLys 81
OY 256 AAGGCTTTTAAAGTGTGTTGACTTGAATGGGGAATGGGGAATGTGAGAGACTTGA 1977
Db 82 GluAlaPheAsnValPheAspGlnAsnGlyAspGlyPheIleThrValAspGluLeuLys 1010
OY 196 TACGTGCTGGGGAAGACTGGGTATGTGAGGTGAAATAAGTGAAGAAAGACTGCAGAGCATG 1377
Db 102 AlaValLeuSerSerIleuGlyLeuLysGlnGlyThrLeuAspPheCysLysLysMet 1270
LY 136 ATTTGGTATTAGCAGACCAATTCAAGCGGATGGTGTGATTTCAGAAATTCAAAACATG 77
Db 122 IleLysLysValAlaAspValAspGlyArgValAsnTyrLysGluPheArgGlnMet 1410
OY 76 ATG 74
Db 142 Met 142
RESULT 5
022845 PRELIMINARY; PRT; 215 AA.
AC 022845: OSFV55;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative calcium binding protein (calmodulin-like MS53) (Putative
DE Ca2+-binding protein).
DE ATG43390 OR T01024.3 OR ATG43390. T1024.3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RC Rounleay S.D., Techudy M.M., Lin X., Ketchum K.A., Crosby M.L.,
RC Brandon R.C., Spriggs T.A., Mason T.M., Kerlavage A.R., Adams M.D.,
RC Somerville C.R., Venter J.C.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RC Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RC Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC MEDLINE=20387008; PubMed=10929106;
RA Kleihow T., Bhalerao R., Breuer F., Umeda M., Salchert K., Koncz C.;
RT "Functional identification of an arabidopsis snf4 ortholog by
RT screening for heterologous multicopy suppressors of snf4 deficiency in
RT yeast."
RL Plant J. 23:115-122(2000).
RN [5]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Bann J., Carninci P., Chen H.,

```



```

QY 136 ATTTGGATTAGACCAATTCAGCGCATGTTGATTTCAGAAATTCATG 77
DB 186 IIMetGlnValAspAlaAspGlyArgValAsnTyrSgluPheLeuGlnMet 205
QY 76 ATG 74
DB 206 Met 206

RESULT 6
Q9ZR02 PRELIMINARY; PRT; 154 AA.
ID Q9ZR02;
AC Q9ZR02;
DT 01-MAY-1999 (T-EMBLrel. 10, Created)
DT 01-MAY-1999 (T-EMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Putative calmodulin.
GN F4C21.22 OR ATAG03290.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
euroside II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
[1]
RN SEQUENCE FROM N.A.
RA Habermann K., de la Bastide M., Huang E.N., Gnoj L., Schutz K.,
RA Preston R., Calma C., Martienssen R., Parnell L.D., Dedhia N.,
RA McCombie W.R.;
RT "Arabidopsis thaliana BAC F4C21 from chromosome IV near 17 cm.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,
RA Preston R.R., Matzko A., Shah R., O'Shaughnessy A., Rodriguez M.,
RA Sheker M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,
RA Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005375; ABD1457.1; -
DR EMBL; AL161496; CAB7814.1; -
DR HSSP; P02593; IFW4.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 4.
DR ProDom; PD000012; EF-hand; 2.
DR SMART; SM00054; EPH; 4.
DR PROSITE; PS00018; EF_HAND; UNKNOWN 4.
SEQUENCE 154 AA; 17590 MW; EEFSE1IB44BD20E CRC64;

Alignment Scores:
Pred. No.: 7 51e-16 Length: 154
Score: 237.50 Matches: 57
Percent Similarity: 55.15% Conservat: 34
Best Local Similarity: 34.55% Mismatches: 51
Query Match: 22.38% Indels: 23
DB: 10 Gaps: 5

US-10-021-323-13 (1-609) x Q9ZR02 (1-154)
QY 538 CTTAGTAGACCGCATTCGACCGATTGAGAAAGCTCGAAGATGAGATGGCTTC 479
DB 1 MetAspSerThrGluLeuAsnArgValPheGlnMetPheAspGlyAspGlyLys 20
QY 478 GTTAGCTGAGAGAGCTGAATTGGTGTCTCCAGAGAAATCGGCTGTCCCAATTCAGCCTT 419
DB 21 IethThrlsYsglulLeuAsnGlnUserPheLysAsnLeuGlyllele---IleProGlu 39
QY 418 GAGGAATTGAGCGCTTAGTG-----GAAAGCAAGATTGGAAGCTTG 377
DB 40 AspgluLeuThrGlnIlelleGlnLysIleAspValAsnGlyAspGlyCysValAspIle 59
QY 376 GATGAATTCTTGTTCTTTATGAATCATCTCGAACCACCTGACATGTGTGACGAA 317

```

```

DB 60 GluGluPheGlyGluLeuTyrIlysrHlleMet-----Val 71
QY 316 GAGGAGAGAGAGAAATTGTCATTCAGCGCGCGGTGAAGAAAGACAGTACCTTGGC 257
DB 72 GluAspGluAspIleVal-----GlyGluGlu-----AspMetLys 83
QY 256 AAGCTTTTAAAGTTTAACTTGAATGGAGATGGGTTGGGGAGATGAGAGAGCTGAA 197
DB 84 GluAlaPheAsnValPheAspArgAsnGlyAspGlyPheIleThrValAspGluLeuLys 103
QY 196 TACGTCGTGGAAGACTGGTATGTGAGGTGAATAATGTAAGAAAGACTGACAGACATG 137
DB 104 AlaValLeuSerSerIleuLeuLysGlnGlyLysThrLeuGluIuLysArgLysMet 123
QY 136 ATTTGGATTAGACCAATTCAGCGCATGTTGATTTCAGAAATTCATG 77
DB 124 IIMetGlnValAspAlaAspGlyArgValAsnTyrSgluPheLeuGlnMet 143
QY 76 ATTTGATTCATCCCGT 62
DB 144 MetLysLysGlyArg 148

RESULT 7
Q9SU00 PRELIMINARY; PRT; 152 AA.
ID Q9SU00;
AC Q9SU00;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Putative calmodulin.
GN T20X18.210 OR AT4G12860.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
euroside II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
[1]
RN SEQUENCE FROM N.A.
RA Bevan M., Peters S.A., van Staveren M., Dirkee W., Stiekema W.,
RA Bancroft I., Mewes H.W., Mayer K.F.X., Schueler C.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Peters S.A., van Staveren M., Dirkee W., Stiekema W., Mewes H.W.,
RA Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL049640; CAB41003.1; -
DR EMBL; AL161535; CAB78328.1; -
DR HSSP; P02593; IFW4.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 4.
DR ProDom; PD000012; EF-hand; 2.
DR SMART; SM00054; EPH; 4.
DR PROSITE; PS00018; EF_HAND; UNKNOWN 4.
SEQUENCE 152 AA; 17182 MW; 0320E53A8B894752 CRC64;

Alignment Scores:
Pred. No.: 1.72e-15 Length: 152
Score: 234.00 Matches: 56
Percent Similarity: 52.80% Conservat: 29
Best Local Similarity: 34.78% Mismatches: 50
Query Match: 22.05% Indels: 26
DB: 10 Gaps: 4

```

US-10-021-323-13 (1-609) x Q9SR7 (1-152)

QY 538 CTTAGTACGACCGCTTGAACCGCTATTGAGAACCTCGACAAGATGAGATGCTTC 479
 DB 1 MetAspArgGlyValLeuSerArgValPheGlnMetPheAspLysAsnGlyAspGlyLys 20
 QY 478 GTTAGCTGAGAGAGCTGTAATGCTGCTCCAGAGAAATGGCTCTCCCAATTCAGCCTT 419
 DB 21 IleAlaLysAsnGlnLeuLysAspPhePheLysSerValGly--IleMetValProGln 39
 QY 418 GAAGATTGAGACCCCTTACTGAGGAAACCATGTTGAC-----TTG 377
 DB 40 AsnGlnIleAsnGlnMetIleAlaLysMetAspValAsnGlyAspGlyAlaMetAspIle 59
 QY 376 GATGAATTTCTTTCTTTATGATCATCTCGAACCCACTGACATGCTGTGACGA 317
 DB 60 AspGlnPheGlySerLeuTyrglnGlnMetVal-----GlnGln 72
 QY 316 GAGAGAGAGAGAGAAATGGTCATTCAAGCGCGCGGTGAAGAGAGACATGACCTTGG 257
 DB 73 LysGlnGlnGlnGln-----AspMetArg 80
 QY 256 AAGGCTTTAAAGTGTGATGATGAGGAGATGGGTGGGATGTGAGAGCTTGA 197
 DB 81 GluAlaPheAspValPheAspGlnAsnGlyAspGlyPheIleThrAspGlnGlnLeuArg 100
 QY 186 TACGTGCTGAGAAAGCTGGGTATGTGAGGTGAAATAGTGAAGAAAGACTGACAGAGCATG 137
 DB 101 SerValLeuAlaSerMetGlyLeuLysGlnGlyArgThrLeuGlnAspGlySerLysMet 120
 QY 136 ATTGGTATTACGACCAATTCAGACGGCATGTTGATTTTCAGAAATTCAAAAATG 77
 DB 121 IleSerLysValAspValAspGlyAspGlyMetValAsnPheLysGlnPheLysGlnMet 140
 QY 76 ATG 74
 DB 141 Met 141

RESULT 8
 Q9SR7 PRELIMINARY; PRT; 153 AA.

AC Q9SR7; 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Putative calmodulin.
 GN F2103.20.
 RA Arabidopsis thaliana (Mouse-ear cress).
 Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 euroids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN NCBI_TaxID=3702;
 RP SEQUENCE FROM N.A.
 RC STRAIN=C.V. COLUMBIA;
 RA Lin X., Kaul S., Town G.D., Benito M., Creasy T.H., Haas B.,
 RA Ronning C.M., Koo H., Fujii C.Y., Uteback T.R., Barnstead M.E.,
 RA Bowman C.L., White O., Nierman W.C., Fraser C.M.,
 RT "Arabidopsis thaliana chromosome III BAC F2103 genomic sequence."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBS databases.
 DR EMBL; AC009853; LFM4.
 DR HSSP; P02593; 1FM4.
 DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; ehand; 4.
 DR ProDom; PD000012; EF-hand; 2.
 DR SMART; SM00054; EFh; 4.
 DR PROSITE; PS00018; EF_HAND; UNKNOWN 4.
 SQ SEQUENCE 153 AA; 17392 MW; ASD5D36363DECAACD CRC64;

Alignment Scores:
 Pred. No.: 9, 116-15
 Score: 227.00
 Percent Similarity: 54.04%

Length: 153
 Matches: 56
 Conservative: 31

Best Local Similarity: 34.788
 Query Match: 21.398
 DB: 10
 Gaps: 4

US-10-021-323-13 (1-609) x Q9SR7 (1-153)

QY 538 CTTAGTACGACCGCTTGAACCGCTATTGAGAACCTCGACAAGATGAGATGCTTC 479
 DB 1 MetAspGlnAlaGlnLeuAlaArgIlePheGlnMetPheAspArgAsnGlyAspGlyLys 20
 QY 478 GTTAGCTGAGAGAGCTGTAATGCTGCTCCAGAGAAATGGCTCTCCCAATTCAGCCTT 419
 DB 21 IleMetLysGlnGlnLeuLysAsnAspSerLeuGlnAsnLeuGly--IleTyrlleProAsp 39
 QY 418 GAAGATTGAGACCCCTTACTGAGGAAACCATGTTGAC-----TTG 377
 DB 40 LysAspLeuValGlnMetIleGlnLysIleAspLeuAsnGlyAspGlyTyrlValAspIle 59
 QY 376 GATGAATTTCTTTCTTTATGATCATCTCGAACCCACTGACATGCTGTGACGA 317
 DB 60 GlnGlnPheGlyGlyLeuTyrglnThrIleMet-----GlnGln 72
 QY 316 GAGAGAGAGAGAGAAATGGTCATTCAAGCGCGCGGTGAAGAGAGACATGACCTTGG 257
 DB 73 ArgAspGlnGlnGln-----AspMetArg 80
 QY 256 AAGGCTTTAAAGTGTGATGATGAGGAGATGGGTGGGATGTGAGAGCTTGA 197
 DB 81 GluAlaPheAsnValPheAspGlnAsnAspGlyPheIleThrValGlnGlnLeuArg 100
 QY 196 TACGTGCTGAGAAAGCTGGGTATGTGAGGTGAAATAGTGAAGAAAGACTGACAGAGCATG 137
 DB 101 SerValLeuAlaSerMetGlyLeuLysGlnGlyArgThrLeuGlnAspGlySerLysMet 120
 QY 136 ATTGGTATTACGACCAATTCAGACGGCATGTTGATTTTCAGAAATTCAAAAATG 77
 DB 121 IleSerLysValAspValAspGlyAspGlyMetValAsnPheLysGlnPheLysGlnMet 140
 QY 76 ATG 74
 DB 141 Met 141

RESULT 9
 Q9F19 PRELIMINARY; PRT; 181 AA.

AC Q9F19; 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Calmodulin-like protein.
 OS Arabidopsis thaliana (Mouse-ear cress).
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 euroids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN NCBI_TaxID=3702;
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=9397451; PubMed=10470850;
 RA Kaneo T., Katoh T., Sato S., Nakamura Y., Asamizu E., Kotani H.,
 RA Miyajima N., Tabata S.,
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. IX.
 RT Sequence features of the regions of 1,011,550 bp covered by seventeen
 RT pl and TAC clones."
 RL DNA Res. 6:183-195(1999).
 DR EMBL; AB017065; BAB09153.1; --
 DR HSSP; P02593; 1CTR.
 DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; ehand; 3.
 DR ProDom; PD000012; EF-hand; 2.
 DR SMART; SM00054; EFh; 3.
 DR PROSITE; PS00018; EF_HAND; UNKNOWN 2.
 SQ SEQUENCE 181 AA; 20253 MW; 7DC8F0E097D57962 CRC64;

Alignment Scores:

Pred. No.:	7.14e-14	Length:	181
Score:	218.50	Matches:	60
Percent Similarity:	50.00%	Conservative:	24
Best Local Similarity:	35.71%	Mismatches:	59
Query Match:	20.59%	Indels:	
DB:	10	Gaps:	5

US-10-021-323-13 (1-609) X Q9F119 (1-181)

```

OY 544 CCCCC-----CTTAAGTAAGACGCACTTGCAACCGGTTGTTGAGAAAGCTCCGACAAAGAAATGGA 488
Db 21 SerProSerLeuAspAlaLeuArgLeuH; sA; gValA.PheaspLeuPheaspLeuAspLeuAsn 40
OY 487 GATGACCTTCGTATGCTCTGAGAGAGCGTAATTGGTCTCTCCAGAGAAATCGAGTGTGTCAA 428
Db 41 AspGlyPheIleIleThrValGlnIleuLeuSerGlnAlaLeuSerArgLeuGly---LeuAsp 59
... 427 TTCAGCCTTGAAGAAATTGAGCCCTTAGTG-----GGAACCATGT 386
... 60 AlaAspPheSerAspLeuLysSerThrValAspSerPheIleLysProAspLysThrGly 79
OY 385 TTGAACCTTGATGAATCTCTGTTCTTTATTAATTCATCTCCGAACCCACTGGACATAGT 326
Db 80 LeuArgPheAspAspPheAlaLeuH; sI; sLysThrLeu----- 92
OY 325 GGTGACGAAGAGAGAGAGAAATTGGTCATTCACGGCGCGGTAA----- 278
Db 93 -----AspIleSerPhePheGlyGlyGlnLysSerCysAspGly 106
OY 277 GAAGAAGACAGNACCTTGCGAAGGCTTTAAAGTCTTGACCTTGATGGATGGGTTG 218
Db 107 SerProGluSerAspLeuGlnIleuAlaPheAsnVal.PheaspLeuAspGlyAspGlyPhe 126
OY 217 GGGGAGATGAGAGACCTTGAAATACGTGCTGGAGAAGCTGGGTATGTAGGTGAAATAGT 158
Db 127 IleserAlaValGlnIleuGlnIleuValIleuLysLeuGlnIleuProGlnAlaGlyGln 166
OY 157 GGAAGAAGACTGCAGAGACATGATTTGGTATTACGACACCAATTCAGCGGCATGTTGAT 98
Db 147 IlegIuGlnValGlnIleuLysMetIleValSerValAspSerAsnHisAspGlyArgValAsp 166
OY 97 TTTCAGAAATCAAAAAACATGATG 74
Db 167 PhePheGlnPheLysAsnMetMet 174
-- "TULP 10
-- R00
-- Q9ER00 PRELIMINARY; PRT; 205 AA.
AC Q9ER00;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE Ayr2/CE-9 rapidly elicited protein 31.
EN ACR31.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; eusteriids I; Solanales; Solanaceae; Nicotiana.
CX NCBL_TaxId=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. PETITE HAVANA;
RA Durrant W.E., Rowland O., Piedras F., Hammond-Kosack K.E.,
RA Jones J.D.G.;
RT "cDNA expression profiling reveals rapid, resistance gene-dependent,
RT active oxygen-independent, gene induction during the plant defense
RT response.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR HSBP; P02593; 1FW4
DR InterPro; IPR002048; EF-hand.

```

DR Pfam; PF00036; ethand; 3.
DR ProDom; PD000012; EF-hand; 2
DB SMART; SM00054; EFh; 3

Alignment Scores:	
Pred. No.:	8.27e-14
Score:	218.00
Percent Similarity:	54.2%
Best Local Similarity:	23.1%
Query Match:	20.5%
DB:	10
Length:	2
Matches:	2
Conservative:	3
Mismatches:	6
Indels:	1
Gaps:	4

US-10-021-323-13 (1-609) x Q9FR00 (1-205)

Qy	544	TCCCCC---CTTAAAGACCGCACTGTGGACGCGTATTCCGAGACGCAAGAAATGGA	488
Db	37	SEPPSESLLEuamSerIleArgLeuNrgArgIlePheAspValPheAspArgAsnHis	56
Qy	487	GATGCGCTTGCTAGTCTGGAGAGCTGAATGTTGCTCCAGAAATCGGCTGTCCAA	428
Qy	427	TTGACGCTTGAAGAAATTGGAGCCCTTAGTg-----GGAAACCATGT	386
Db	76	AlaAspGlnSerGlnIleGlnSerMetValArgSerTyrIleTylsSerGlyAsnAngly	95
Qy	385	TTGAACCTGGAGAAATTCCTGTCTTTATATACCATCTCGAACCCACATGCAATGCT	326
Db	96	LeuArgPheGlnAspPheGlnAlaLeuHisIleArgSerLeuAspArg-----ValPhePhe	113
Qy	325	GGTGAAGAAAGAGAGAGAGAAATGTGCTATTCACGCGCGGTGAAGAAGACAGT	266
Db	114	GlySerTyrTylGlnGlnAspArgTyrIleValLeuAspGlnAspProAspGlnAspGlnVal	133
Qy	265	GACCTTGCGAAGGCTTTTAAAGTCTTTTACTGTAATGGAGATGGCTTGGGGGAGTGGAG	206
Db	134	AspLeuLysAspAlaPheAsnValPheAspGlnAsnGlyAspGlyPheIleSerAlaLys	153
Qy	205	GACCTTGAATPAGTGTGCTGGAGACACTGCGTATGTGAGGTGAATAATGTGAAAAAGACTGC	146
Db	154	GlnLeuGlnAlaValLeuGlnLysLeuGlyLeuProGlnGlyAsnGlnIleAspArgVal	173
Qy	145	AGGACGATGATTTGGTATTACGACCAATTCAGACGCGATGCTGTATTTTCAAGAAATTC	86
Db	174	GlnMetMetIleSerSerValAspGlnAspHisAspGlyGlnValAspPheValGlnPhe	193
Qy	85	AAAAACATGATG 74	
Db	194	LysAspMetMet 197	
RESULT 11			
Q9LXZ7			
Q9LXZ7	PRELIMINARY;	PRT;	195 AA.
AC	Q9LXZ7;		
DT	01-OCT-2000 (T:EMBLrel. 15, Created)		
DT	01-OCT-2000 (T:EMBLrel. 15, last sequence update)		
DT	01-JUN-2002 (T:EMBLrel. 21, last annotation update)		
DE	Calmodulin-like protein.		
GN	F25L23_300.		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsids.		
OK	NCBI_TaxID=3702;		
RN	(1)		
RP	SEQUENCE FROM N.A.		
RA	D'Angelo M., Vezzi A., Modesto D., Pigazzi M., Valje G., Mewes H.W.,		
RA	Rudd S., Lemcke K., Mayer K.F.X., Quezzer F., Salanoubat M.;		
RL	Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.		
RF	(2)		
RF	SEQUENCE FROM N.A.		

Db 138 PheValAsnIleIle 142

RESULT 15

Q9ATP2

PRELIMINARY; PRT; 188 AA.

AC Q9ATP2; 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 OS Calmodulin-like protein.
 OS Penisetum ciliare (Buffelgrass).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoideae; Paniceae; Penisetum.
 OX NCBI_TaxID=35520;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PISTIL;
 RA Li Z., Burson B.L., Zhang F., Hussey M.A.;
 "Isolation and Characterization of Genes Differentially Expressed in
 the Pistils of Apomictic Buffelgrass (Penisetum ciliare [L.] Link.)";
 Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF325719; AKI5501.1; -.
 DR HSBP; P02593; ICTR.
 DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; ehand; 3.
 DR ProDom; PD000012; EF-hand; 2.
 DR SMART; SM00054; Eph; 3.
 DR PROSITE; PS00018; EF_HAND; UNKNOWN 3.
 SQ SEQUENCE 188 AA; 20250 MW; 90135IDA38EB1CB3 CRC64;

Alignment Scores:

Pred. No.:	4.1e-12	Length:	188
Score:	201.50	Matches:	56
Percent Similarity:	52.80%	Conservative:	29
Best Local Similarity:	34.78%	Mismatches:	61
Query Match:	18.9%	Indels:	15
DB:	10	Gaps:	5

US-10-021-323-13 (1-609) x Q9ATP2 (1-188)

QY 538 CTTAGTAAGACCGACTGCAAGCCGATTTCGAGAGACTCGACAAGATGAGATGCTTC 479
 |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
 Db 29 LeuAspAlaLeuArgLeuArgAlaValPheAspLeuPheAspArgAsnGlyAspGlyGlu 48
 478 GTTAGCTGGAGAGCTGAATGCTGCTCCAGAGAACTGGGCTGTCGCAATTGACGCTT 419
 :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
 49 IleThrLeuAspIleuMetAlaSerAlaLeuAspThrLeuGly--LeuGlyAlaAspArg 67
 418 GAGAAATGAGAGCCCTTAGTGGGAAA-----CCATGT-----TTGAACCTG 377
 |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
 Db 68 SerGlyLeuGluAlaAlaValGlySerTyrIleProAlaGlyAlaIleGlyLeuArgPhe 87
 QY 376 GATGAATTTCTGTTCTTTATGATCATCTCGAACCCACTGACACATGCTGTCGAA 317
 :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
 Db 88 GluAspPheGluSerLeuHisArgAlaLeuGlyAspAlaLeu--PheGlyProIlePro 106
 316 GAGGAGAGAGAGAAATGTCATTCACGGCGCGGTGAGAGAGACAGTACCTTGGC 257
 |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
 Db 107 GluGluValProGlu-----GluAspAspGluGlyAspMetLys 119
 QY 256 AAGGCTTTTAAAGTGTGATCTGAATGGGATGGGCTGGGATGAGAGCTTGA 197
 :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
 Db 120 GluAlaPheArgValPheAspGluAspGlyAspGlyTyrIleSerAlaIleGluLeuGln 139
 QY 196 TACGTCTGGAGAGACTGGGTATGTGAGGTGAATAATAGTGAAGAAGCTGACAGACATG 137
 |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
 Db 140 AlaValLeuLysLysLeuGlyLeuProGluAlaArgAsnLeuAlaThrValGlnGluMet 159
 QY 136 ATTTGGATTATGACACCAATTCAGACGGCATGTTGATTTCAGAAATTCATAAATCATG 77
 |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
 Db 160 IleCysAsnValAspAlaAspArgAspGlyArgValAspPheGlyGluPheLysAsnMet 179

QY 76 ATG 74
 |||
 Db 180 Met 180

Search completed: June 24, 2003, 14:05:08
 Job time : 52 secs

Mon Jun 30 09:38:05 2003

us-10-021-323-13.rge

MARTIN
10/02/2003 Page 1
209.10 B

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2003, 06:12:44 ; Search time 1245 Seconds

(without alignments)
14235.830 Million cell updates/sec

Title: US-10-021-323-13

Perfect score: 609

Sequence: 1 ggaatgaatcaacttc.....tgaagcctacaactaag 609

Scoring table: IDENTITY NUC

Searched: 2054640 seqs, 1453402878 residues

1 number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

GenEmbl: 1: gb ba: 2: gb hcg: 3: gb in: 4: gb om: 5: gb ov: 6: gb pat: 7: gb pl: 8: gb pr: 9: gb ro: 10: gb scs: 11: gb sy: 12: gb un: 13: gb vi: 14: gb vl: 15: em ba: 16: em fun: 17: em hum: 18: em in: 19: em mu: 20: em om: 21: em or: 22: em ov: 23: em pat: 24: em pl: 25: em ro: 26: em scs: 27: em un: 28: em vl: 29: em vl: 30: em hcg hum: 31: em hcg inv: 32: em hcg other: 33: em hcg mus: 34: em hcg pin: 35: em hcg rod: 36: em hcg mam: 37: em hcg vrt: 38: em sy: 39: em hcg hum: 40: em hcg mus: 41: em hcg other:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	157	25.8	7758	8	AP004923
2	136.2	22.4	80229	2	AC130200
3	102.2	16.8	777	8	AF361594
4	102.2	16.8	108365	8	AC015447
5	96.8	15.9	429	8	AF133593
6	77.4	12.7	97076	8	H0522A01
7	75.8	12.4	88560	8	OSJN00138
8	75.8	12.4	133868	8	OSJN00061
9	70	11.5	79375	8	ATT20K18
10	70	11.5	199280	8	ATCHRIV35
11	65.2	10.7	159465	2	AP005001
12	63.2	10.5	90112	8	ATAC009853
13	63.2	10.4	173645	2	AC123526
14	62.2	10.2	631	8	LJA251808
15	60.2	9.9	119048	2	AP004167
16	60.2	9.9	148076	2	CNS08C98
17	57	9.4	159191	2	AP005099
18	55.6	9.1	69563	8	ATF21C20
19	55.6	9.1	126957	2	AC122169
20	55.6	9.1	197568	8	ATCHRIV53
21	53.6	8.8	778	8	AY081574
22	53.6	8.8	907	8	AF250344
23	53.6	8.8	972	8	AY087087
24	53.6	8.8	993	8	AY062815
25	53.6	8.8	90283	8	AC004450
26	53.6	8.8	103057	8	AC002335
27	52.4	8.6	108992	8	AP003740
28	52.4	8.6	142081	8	AP004572
29	52.2	8.6	453	8	AF332466
30	52.2	8.6	74316	8	T21E18
31	52.2	8.6	95600	2	AC024227
32	52.2	8.6	152170	2	AP005506
33	51.6	8.5	7218	6	I66494
34	51.2	8.4	143831	2	AP005526
35	49.8	8.2	975	8	AF211529
36	49.4	8.1	570	8	MSA295608
37	48.2	7.9	211657	2	AL732564
38	48	7.9	61290	8	AB017065
39	47	7.7	762	8	BVBEV111
40	46.6	7.7	205015	10	AC007937
41	46	7.6	186516	2	AC116882
42	45.6	7.5	180246	10	AL627186
43	45.6	7.5	198516	2	AC101761
44	44.6	7.3	204495	2	AL671893
45	44.6	7.3	279124	2	AL672268

ALIGNMENTS

RESULT 1
AP004923/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE

AP004923 73758 bp DNA linear PLN 19-JUL-2002
Lotus japonicus genomic DNA, chromosome 1, clone: LJT18013, TM0078,
complete sequence.
AP004923.1 GI:21907939
HTG.
Lotus japonicus DNA, clone: LJT library clone: LJT18013.
Lotus japonicus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
Lotus.

RESULT 3
AF361594/C

16.8%; Score 102.2; DB 8; Length 777;

Best Local Similarity 53.5%; Pred. No. 1.8e-17;
Matches 255; Conservative 0; Mismatches 198; Indels 24; Gaps 1;

```

QY 68 ATGTATACATCATGTTTGAATCTTGAATAACACGCGCTGAAATGGTGCAGA 127
Db 554 ATGTATACATCATGTTTGAATCTTGAATAACACGCGCTGAAATGGTGCAGA 495
QY 128 ATACCAATCATGCTCTGAGCTTCTTCCATATTTTCCACATACCCAGCTTCC 187
Db 494 AACTGCAATCATGCTCTGAGCTTCTTCCATATTTTCCACATACCCAGCTTCC 435
QY 188 CAGCAGATTCAGTCTCTGACATCCCCCAACCATCCCATTCAGTCAACACTTT 247
Db 434 CACACATCATCAGAGCTCTCTGAGCTTCTTCCATATTTTCCACATACCCAGCTTCC 375
QY 248 AAAAGCTTCGAGAGTCACTGCTCTTCTTCCACGCGCGCTGATGACCAATTCCTC 307
Db 374 AAAAGCTTCGAGAGTCACTGCTCTTCTTCCACGCGCGCTGATGACCAATTCCTC 315
QY 308 CTCCTCTCTCTGCTGACCAATGCTGAGCTGCTGAGATGATTCATTAAGAACA 367
Db 314 TTTCATATCTAAGAGC-----CTCGATGAGAACGGAAG 279
QY 368 GAATTCATCCAACTTCAACATGCTTTTCCCACTAAGGCTCCAAATTTTCAAGCTGAA 427
Db 278 GAATTCATCCAACTTCAACATGCTTTTCCCACTAAGGCTCCAAATTTTCAAGCTGAA 219
QY 428 TTGGAAGACAGCCGATCTCTGAGAGCAACCAATTCAGTCTCTGAGATGATTCATTAAGAACA 487
Db 218 TTGGAAGACAGCCGATCTCTGAGAGCAACCAATTCAGTCTCTGAGATGATTCATTAAGAACA 159
QY 488 TCCATCTCTGCTGAGCTTCTGAGATGCTGAGCTGCTGAGATGATTCATTAAGAACA 544
Db 158 TCCATCTCTGCTGAGCTTCTGAGATGCTGAGCTGCTGAGATGATTCATTAAGAACA 102

```

RESULT 4
AC015447 108365 bp DNA linear PLN 26-JUL-2000
LOCUS Arabidopsis thaliana.chromosome 1 BAC F2408 genomic sequence,
DEFINITION complete sequence.
AC015447
VERSION AC015447.8 GI:7340330
KEYWORDS HTG.
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana.

REFERENCE
AUTHORS
1. (bases 1 to 108365)
Federpspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F.,
Altafi, H., Nguyen, M., Lam, B., Southwick, A., Bel, O., Gonzalez, A.,
Buenher, E., Chou, J., Choi, E., Chou, J., Choi, E., Gonzalez, A.,
Hewling, B., Johnson-Hopson, C., Khan, S., Kim, C., Koo, T., Lee, J.M.,
Leng, C., Liu, A., Liu, S., Mukharzky, N., Pham, P., Sakano, H.,
Shim, P., Toriumi, M., Vaysberg, M., Yu, G., Ecker, J., Theologis, A.
and Davis, R.W.

JOURNAL
REFERENCE
AUTHORS
2. (bases 1 to 108365)
Federpspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F.,
Altafi, H., Nguyen, M., Lam, B., Southwick, A., Bel, O., Gonzalez, A.,
Buenher, E., Chou, J., Choi, E., Chou, J., Choi, E., Gonzalez, A.,
Hewling, B., Johnson-Hopson, C., Khan, S., Kim, C., Koo, T., Lee, J.M.,
Leng, C., Liu, A., Liu, S., Mukharzky, N., Pham, P., Sakano, H.,
Shim, P., Toriumi, M., Vaysberg, M., Yu, G., Ecker, J., Theologis, A.
and Davis, R.W.

TITLE
JOURNAL
Direct Submission
Submitted (16-NOV-1999) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA (bases 1 to 108365)
3. (bases 1 to 108365)
Federpspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F.,
Altafi, H., Nguyen, M., Lam, B., Southwick, A., Bel, O., Gonzalez, A.,
Buenher, E., Chou, J., Choi, E., Chou, J., Choi, E., Gonzalez, A.,
Hewling, B., Johnson-Hopson, C., Khan, S., Kim, C., Koo, T., Lee, J.M.,
Leng, C., Liu, A., Liu, S., Mukharzky, N., Pham, P., Sakano, H.,
Shim, P., Toriumi, M., Vaysberg, M., Yu, G., Ecker, J., Theologis, A.
and Davis, R.W.

Chin, C., Chiu, J., Choi, E., Dunn, P., Gonzalez, A., Hong, B., Kim, C.,
Koo, T., Lee, J.M., Leng, C., Li, J., Liu, A., Liu, K., Liu, S.,
Mukharzky, N., Pham, P., Sakano, H., Schwartz, J., Shim, P.,
Thaveri, A., Toriumi, M., Vaysberg, M., Walker, M., Yu, G., Ecker, J.,
Theologis, A. and Davis, R.W.
Direct Submission
Submitted (29-MAR-2000) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
4. (bases 1 to 108365)
Federpspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F.,
Altafi, H., Nguyen, M., Lam, B., Southwick, A., Ecker, J., Theologis, A.
and Davis, R.W.
Direct Submission
Submitted (26-JUN-2000) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
On Mar 29, 2000 this sequence version replaced gi:7021718.
Bases 98,795-108,365 of clone F2408 overlap with bases 1-9,571 of
IGF BAC clone F8X7, gb|AC007727.
Location/Qualifiers
1. 108365
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="1"
/clone="F2408"
complement(5411..7047)
/gene="F2408.1"
complement(join(5411..5453,5533..5598,5673..5755,
6043..6105,6199..6267,6350..6403,6477..6557,6652..6717,
6814..7047))
/gene="F2408.1"
/note="Hypothetical protein; similar to hypothetical
protein Rv1732c from Mycobacterium tuberculosis,
Prl|G70687"
/codon_start=1
/protein_id="AAF87891.1"
/db_xref="GI:9454568"
/translation="MASTVVTVTTHGASILPRTALAVSRRFSFSPRLDPSNL
RRNVAQSPRRRLVVARTEASAGVLAGAPPELPETLGNLMLDEPLPSLLM
PICHNCPYTHAKDIYVLCNPKMKGLAVAISSNVTHHQDGPETPAKAPKPY
PPVLYDSQVAREFAVCTPEFLYKDGKRPPELVYHGGFDSPSSNSPVTGRD
LSLAIDLISCPISNOKPSVGSIKMPEKTS"
10192..11472
/gene="F2408.2"
join(10192..10301,10635..10770,10948..11091,11238..11472)
/gene="F2408.2"
join(10639..10770,10948..11091,11238..11273)
/gene="F2408.2"
/note="Unknown protein"
/codon_start=1
/protein_id="AAP87892.1"
/db_xref="GI:9454569"
/translation="MDFTAGLRLLSKENSSKMEELVEBCYTTTKPKHGWITAAAFK
VTKVLPDKMTWEAIGARDESDYDLREDIDTSLTPILKEIYFVLEQGLSLRLS
M"
complement(13842..16068)
/gene="F2408.3"
complement(join(13842..14451,14532..14720,14817..14902,
14985..15066,15162..15313,15396..15509,15599..15697,
15780..15858,15959..16068))
/gene="F2408.3"
/note="Unknown protein"
/codon_start=1
/protein_id="AAF87893.1"
/db_xref="GI:9454570"
/translation="MANNAAACARATNDMLIGPMAINIELCDIINNEPQAKAVK
VKKRIGSKNSKYQVILATVLETLKSCESYQULIVDDILPDKVTKKPPDLTV
KTLISLDTWQAFSGGGRFOYVNAVYELRSGAGLEFPPTSSVFPPTPOTPIV
AORTASDENALIOALISQSDPASLMEELIQSQSGVADVLTDLGALDPEHPGSEEL
IVDLVQECTTQRRVWALVNTTSDELMKQGLALNDLQRYVQHDDDAKKSQVAPATA
PPPIPLVTSINHDDESDDDFLQALHRSKRSAGTGGGNFNPPLPPSSMRPVHV
DSGAMDFLSGDVYKQETPEENKPPSTSSSNHDYAPLIFDEPPVQSSPEHALFTKP

COMMENT
FEATURES
source

gene
CDS

gene
mRNA
CDS

gene
CDS

gene
CDS
21528..24276
/gene="F2408.4"
join(21528..21761,22480..22617,22699..22861,22959..23116,23203..23302,23402..23513,23608..23747,23837..24022,24109..24276)
/gene="F2408.4"
/note="Nearly identical to branched-chain alpha keto-acid dehydrogenase E1 - alpha subunit, gb|AAC69851"
/codon_start=1
/product="branched-chain alpha keto-acid dehydrogenase E1 - alpha subunit"
/protein_id="AAF87894.1"
/db_xref="GI:9454571"
/translation="MAIWFAFSKTLVSSLRHNLIITLLIKRDYSHRPIFYTTQSSS
TAVLSFPGSLHSESTAVETQADHLVQOIDEVDAOELDPGGVKGTSKMKPESSSR
RIPCVRLEDGRIPDSDFPVSERLAVMEQWATQVMDHIEYEAQORGLISFL
TSGBEAININISAAALSPDDVLPQREPVLMRGPFTLEEFANCPGNKADYGRRO
MTHGSRNLVYFTSSRIATQVLPQAGVGLKDKKACATPFIQGGGTSSEDFRA
GLNFAVMEARVFIQRNGWALSTHISEQFSDGVYKQAYGIRSRVNDNDLAV
YSAVSAKEMAVTEQRPVLIEMTYRVGHSTSDSTYRADEIQYKMSNPVNR
KRWEDNGMSEDESKLSNARKQLQAIQAEMKQPLTELFNDVYDVYKPLIE
DELKELVKOPDYPGGFHY"
25515..27422
/gene="F2408.5"
join(25515..25732,26281..26629,26734..26875,27049..27422)
/gene="F2408.5"
/note="Unknown protein"
/codon_start=1
/protein_id="AAF87895.1"
/db_xref="GI:9454572"
/translation="VWGEASMEIDQCFQKMEGSIKEMKDIPEVLEMLRISLV
DRNYSVAGVCTGMRDAISFGLRLRLSCNNNNNSLVLSPKRVKIQTLNRDCK
OLEDNAVAIAHNCHELOELDSKLTDSLALAGCPDLTALNSGTSFSDTA
IAYLRFQRKLVNLGCGKAVTNALAEIGNCNOGMSLIGCEMISDGVSLA
YCPDLRTLDLCCGLITDESVALADMCVRLSLGLYCRNITPRAVYSLAOSGVK
KPGSKSVKKGKYEDEGLRSNLISCTLTLSAVQAVCDSPALHTCGRSHLVNSGC
LNLTHCACTLQAHRAHNAVPHAH"
26881..30222
/gene="F2408.6"
join(26881..29498,29585..29941,30022..30222)
/gene="F2408.6"
/note="Similar to flavin-containing monooxygenases"
/codon_start=1
/protein_id="AAF87896.1"
/db_xref="GI:9454573"
/translation="MEKIKILITIGAPAGLATACINRLNINIVRVPVCSAL
WKRSDYDLKHLAKQFCOLPMPFPSTPTPTVSLGRIINTIDEVATFNNVRNRL
VKSAYFKQGQIVKVNKTALILEVSAKFWAAAGENGEGVPEIPEGLVESFOGKIL
HSSEYKNEKFAKQVLLVCGNSGMEIAYLDSKNANVSIYRQVHLTRCIVAG
MSLLEFPVKLVDRCLLLELRFNRTSGVLRPNNGFLAKLTGSAITLDCVAG
EIKSGKIOVTSIKRIEKGTVFIDGNTKVDISIVFAQYKSVKMLSEVDGDFNE
NGMPKREPRDHKNGKGLYSAGFGKQGLAGISRDARNLARDIDSLVCGRSSKNKLSK"
complement(30361..32139)
/gene="F2408.7"
complement(join(30361..30726,30817..30972,31421..31503,31740..31772,31863..32139))
/gene="F2408.7"
/note="Similar carboxyphosphonenoelpyruvate phosphonmutases"
/codon_start=1
/protein_id="AAF87897.1"
/db_xref="GI:9454574"
/translation="MSMLAAKSTLSLFSNPTISAKIGONPRGVSVYPTVMQSRV
RLIEOGAVLIGVVDALSAIVQOTGSAAALISGVALSALTGDPDDILMOORRD
CYTGCGNLNORTYKDLIAGACGFLVPAEBAKAKISARALIDADFLIART
DARLASATGSDAIDRANLWEGADASFEVAPDDDELKEIGRTGKYLGNLWEG
GRTPLATDELKEMGFHLIAHLELTSYSTRALVVLKLIKKEITKHOLEXMTTFEE
FNRLVLDWSVLETKYSNRLALGETK"
31772..33936
/gene="F2408.8"
/note="SCL1"

Query Match 16.8% Score 102.2; DB 8; Length 108365;
Beet Local Similarity 53.5%; Pred. No. 3.9e-17;
Matches 255; Conservative 0; Mismatches 198; Indels 24; Gaps 1;
rnna
join(37172..37211,37747..39836)
/gene="F2408.8"
3757..39538
/gene="F2408.8"
/note="scarecrow-like 1 protein; member of the GRAS protein family, SCL1, cDNA clone gb|AF210731"
/codon_start=1
/protein_id="AAF87898.1"
/db_xref="GI:9454575"
/translation="MVEQTVREIKRWSLVNSAEPSSYRPNKLTLLNNGNNGV
SSAQIFPDERKNPCLTDDSYQSXYEFLDSPTDEVQHPISGASVSSFGSLDF
PYQRPVYGCSEMEQPLDSTSTSTRIGQAVSYSPMDVVEEPDEQMRKIOE
LERALGVEDDQWQVNLNMDISSEYSONSEDOHSPKSSSADSNVSHKSVYS
QATPKOILISCARALSGKLEALSMNLEROIYISIGDPSORIAAYVVEGLAARMA
SGRTYRALKKEKPEPSDBRLAMQVYLVCCCFRFGELAAAGALIEAIKEEVEHID
FDIQGQVMTLISIALBPKRRLRLTGIDDESQVRSIGGIRIGRLRQLAEDN
GVSFRFAMPKSTIVSPSTLGCRRGELLYNFAFOLHMDSESVTVNQDELHNV
KSLNPLKTVTEQVNTNTPFPFRFEAYEYGAVESLDMTLPREQEMNVERQC
LARDIVINACEGERIEREYBAKMRAMMAGFNPKMSAKVTNNIQLIKOOYCN
68 ATGTACATCATGTTTGAATTCCTGAAATCAACCAATGCGCTGATTTGTCGTA
81359 ATGTAGATCATGTTCTTGAATTTCTTCAAGTCAACAAACATCAATTTCTTCATG
128 ATACCAATCATGCTCTCTGAGTCTTTTCCACTATTTTCACTCATATCCAGTCTTC
81419 AACTGATCATTTCTCCCAATCCAGACCTTTGCTCCCTCCCAACCTGACCGTTC
188 CAGCAGTATTCAGTCTCTCAATCCCCCAACCCATCCCATTCAGTCAACACTTT
81479 CAACACATCAAGAAAGCTCTCCGCTGAATATTAACCATCTCGTTCATGAAACACTT
248 AAAAGCCTTCCGCAAGGTGACGTCTCTCTTCCACCGCGCGGGAATGACCAATCTCTC
81539 AAAAGCCTTCCGCAATCGGTTCAATATTAACGAAACCAATATTTCTTTAGATCC
308 CTCCTCTCTTCGTCACCAACCATGTCAGTGGGTTCCAGATGATTCATTAAGAACAA
81539 TTGCTACTTAACACAGC-----GTGCTGTGAACGGAAG 81634
368 GAATTCATCAAGTCAACATGTTTTCACCACTAAGGCTTCAATTTTCAAGCTGAA
81635 GAATCGCTTAATCAAGAGCTCTGTTTCCACAGATCACTTCAAGCTATCGGGGTG
428 TTGACACACCGGATCTCTGAGACCAACATTCAGCTCCCGACATCAAGAGCCATC
81695 TTGGGCCCCAGGATTATCAAGATCAAGAGGCTGTCGAGGGTCACAAAGCCGCTC
488 TCCATTTCTTTCGAGCTTCTGAAATACGCGTTGCAAGTGGGCTTACTAAGGGGGA
81755 TTGGTTCTTTCGAGCGCTTGAACATCTTTCGAGATCTTTGAGTTAAGATGA
RESULT 5
AY133593/c 429 bp. mRNA linear. PLN 07-AUG-2002
AY133593 Arabidopsis thaliana At1g21550/F2408_7 mRNA, complete cds.
DEFINITION
ACCESSION AY133593
VERSION AY133593.1 GI:22137155
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi
1 (bases 1 to 429)
Chen, R., Chen, H., Kim, C.J., Shim, P., Ban, J., Bowers, L.,
Caminici, P., Chang, E., Dale, J.M., Goldenhith, A.D., Hayashizaki, Y.,
Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawat, J.,

TITLE
JOURNAL
AUTHORS

Arabidopsis ORF clones
Unpublished
2 (bases 1 to 429)

Check, R., Chen, H., Kim, C.J., Shin, P., Ban, J., Bowser, L.,
Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,
Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,
Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,
Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K.,
Yamamura, Y., Yu, G., Yu, S., Shinzaki, K., Davis, R.W., Theologis, A.,
and Ecker, J.R.

Direct Submission
Submitted (17-JUL-2002) Salk Institute Genomic Analysis Laboratory
(Signal), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA

RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RFLP CDNA (RFLP CDNA: 'RIKEN
Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y., and Shinzaki, K.

The Salk, Stanford, RGC (SSP) Consortium members constructed and
sequenced the pPUN1 (ORF) clones using the RFLP CDNA: Check, R.,
Chen, H., Kim, C.J., Shin, P., Ban, J., Bowser, L., Chan, M.W.,
Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Jones, T.,
Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M.,
Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C.,
Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Davis, R.W.,
Theologis, A., and Ecker, J.R.

Check, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
this work. Shinzaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
contributed equally to this work as PIs.

FEATURES
source
1..429
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="1"
/clone="U12085"
/note="This clone is in pPUN1 51
ecotype: Columbia"
1..429
/note="unknown protein"
/codon_start=1
/product="At1g21550/F24J8.7"
/protein_id="AA091423.1"
/db_xref="GI:22137156"
/translation="MFKTLDKNQDGLVTLIDELMLIDKGMABHTPDELELVKQSL
DIDELRIFYDAVLDSKSGKNLIDVADDEDAIARAFNVDFGQYISAEELRDYLE
RLGFEEAKAMDGCMIRVHDKLIDDFVFEFKNMILHV"

BASE COUNT
114 a 70 c 122 g 123 t

ORIGIN

Query Match 15.9%; Score 96.8; DB 8; Length 429;
Best Local Similarity 53.8%; Pred. No. 5,2e-16;
Matches 240; Conservative 0; Mismatches 182; Indels 24; Gaps 1;

QY 68 ATGTAACATCATCTTTGAATCTTGAATCAACCAACCGCTCGAATGGTCTGTA 127
DB 423 ATGTAGATCATGTTCTTAATTTCTCAAGTCACAAACCAATCAAGATTTCTTCATG 364
QY 128 ATCAACATCATCTCTGAGCTTTTCACTATTTTCAACCATCATCCGAGTCTCC 187
DB 363 AACTGATCATCTCTCCCAATCCCAAGCTTTGCTCTCTCAAAAGCTTACGCTTC 304
QY 188 GAGCAGTATTCAGCTCTCATATCCCAACCAACCATCCCATTCAGTCAACACTT 247

DB 303 CAACATCAAGAAAGCTCTCCGCTGAAATATTAACATCTCCGTCACATGAAACAT 244
QY 248 AAAAGCCCTCCCAAGGTCATGCTCTTCTTCTTCAACCGCCGCGAATGACCAATTCCTC 307
DB 243 AAAAGCCCTCCCAAGGTCATGCTCTTCTTCTTCAACCGCCGCGAATGACCAATTCCTC 184
QY 308 CT 367
DB 183 TTGTGATCTTAAACAGC-----GTCTGTGTGAACCGAAG 148
QY 368 GAATTCATCCAGTTCATCAACATGTTTCCCACTAAGGCTCCATTTCTTAAGGCTGA 427
DB 147 GAACCTGCTTAAATGAGAGCTCTGTTTACCAACGATCTCAAGCTTATCGGAGTGTG 88
QY 428 TTGACAGACCCGATCTCTGAGACCAATTCAGCTCTCTCAAGCTTCAAGCAAGCCATC 487
DB 87 TTGGCCCAACCGAGTTTATCAAGATCCAAAGAGCTGTGAGGCTCACAGCCGCTC 28
QY 488 TCCATCTTTGTCAGCTTCTTGATA 513
DB 27 TTGCTTCTTGTGAGCGCTTGAAACA 2

RESULT 6
H0522A01 97076 bp DNA linear PLN 08-JAN-2001
LOCUS
DEFINITION
Oryza sativa genomic DNA, chromosome 4, BAC clone: H0522A01,
complete sequence.
ACCESSION
AL512542.1 GI:12140339
VERSION
KEYWORDS
SOURCE
ORGANISM
Oryza sativa.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 97076)
REFERENCE
Han, B., Peng, Q., Mu, J., Zhou, B., Chen, Z. H., Li, Y., Zhu, J. J.,
Tang, Y. S., Zhao, Q., Liu, Y. L., Huang, Y. C., Yu, Z., Fan, D. L., Chen, L.,
Weng, Q. J., Zhang, L., Lu, Y. Q., Yu, S. L., Zhu, X. H., Hu, X.,
Lei, H. Y., Zhang, Y. J., Wang, R., Li, C., Lu, Y., Chen, X. C., Zhang, Y.,
Hu, H., Jia, P. X., Li, T., Qian, Y. M., Jing, K., and Hong, G. F.
Oryza sativa indica (Guangluai4) genomic DNA, chromosome 4, BAC
clone: H0522A01
Unpublished
2 (bases 1 to 97076)
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
HONG, G. F.
Direct Submission
Submitted (03-JAN-2001) Hong G. F., National Center for Gene
Research, Chinese Academy of Sciences, 500# Caobao Road, Shanghai
200233, CHINA

FEATURES
source
1..97076
/organism="Oryza sativa"
/variety="Guangluai4"
/sub_species="indica"
/db_xref="taxon:4530"
/chromosome="4"
/clone="H0522A01"
/clone_11b="BAC-HindIII"

BASE COUNT
26887 a 21448 c 21030 g 27711 t

ORIGIN

Query Match 12.7%; Score 77.4; DB 8; Length 97076;
Best Local Similarity 49.7%; Pred. No. 3.2e-10;
Matches 226; Conservative 0; Mismatches 226; Indels 3; Gaps 1;

QY 74 CATCATGTTTGAATCTTGAATCAACATGCGCTGTAATGCTGTAATACCA 133
DB 80144 CATCATGAGAGAACTGTGGAATGACGCGCGCTGCGCGCTGAGACTGAGC 80203
QY 134 AATCATGCTCTGAGCTTTTTCATATTTTCACTCATACCAATCCAGTCTTCCAGAC 193

	RESULT
OSJN00138/c	7
LOCUS	OSJN00138/c
DEFINITION	OsJN00138 88560 bp DNA linear PLN 17-JUL-2002
	Oryza sativa genomic DNA, chromosome 4, BAC clone: cJ590528_30,
	complete sequence.

REFERENCE
AUTHORS

TITLE
JOURNAL

REMARK	Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC clone: cf990528_30.
COMMENT	On Jul 19, 2002 this sequence version replaced gl:17998450.

Assembly program: phrap

NOTE: This is a complete sequence.

Genes were identified by a combination of several methods: gene prediction programs including Fgsearch (<http://www.softberry.com/>), GENSCAN (<http://CCR-081.mit.edu/GENSCAN.html>), GENSCAN-KMH (<http://genemark.biology.gatech.edu/Genemark/>), tRNAscan-SE (<http://genome.wustl.edu/eddy/tRNAscan-SE/>), searches of the complete sequence against NCBI non-redundant protein database (<http://nbl.nlm.nih.gov/blast/db>) and the EST database at NCGR.

FEATURES
source

Location/Qualifiers
1. .88560

```

/organism="Oryza sativa"
/variant="Niponbare"
/sub_species="japonica"
/db_xref="taxon:4530"
/chromosome="4"
/cione="oj990528_30"
/cione_1bp="Monanto-Ot"
BASE COUNT      2443 a 1846 t 1990 g 2656 c
ORIGIN

```

Query Match	12.4%	Score 75.8 ;	DB 8 ;	Length 88560 ;
Best Local Similarity	49.5%	Pred. No. 8.7e-10 ;		
Matches 225 ;	Conservative 0 ;	Mismatches 227 ;	Indels 3 ;	Gaps 1 ;

74 CATCATGTTTGAATTCCTGAAAAACACCATGCGGCTGAAATGGATCGAATACCA 133

134 AATCATGCTCTGCAGCTTTTCCACTATTTTCACTCATATACCAGTCTTCCAGAC 193

0y 194 GATATTCAAGCTCTCATCATCCCCCAACCATATCCCATTAAGTAAACATTTTAAAGC 253
Db 71451 GGGCCCAAGCTCTCATCATCCCCGATGAAACCGTGGCGTTGGCGTGAACACCCGAAACG 71392

D_b

Oy	254	CTTCGAAGGCACTGTCTTCTCTTACCGCCGCCTGAATGACAATTCTCCTCTCC	313
Db	71391	CTCCCATGTCGGCGCTCTCTCTCTCTCTCTCTTGAGCGCGGCGCTCTGTCTGTC	71333

314 CTCTTCGTCACCACCATGTGGCAGTGGGTTGAGATGATTCATTAAGAACAAGAAATTC 373

374 ATCCAGTCAACATGGTTTCCCACTAAGGGCTCCATTCTTCAAGGCTGAATTGNC 433

434 AGACCGCATTTCTGAGCAACCAATTCAGCTCTCCAGACTAAGCAAGCCATCTCCATT 493

Db 71212 -TCCGAGCTTCCCCAGCAGTCTCTCCAGCTCTCCCGCGGATCCGCCCGCCGCTT 71155

Ov 494 CTTGTGAGCTTCTCGAATACGGCTTGCAAGTCGG 528

Db 71154 CCTGTCGAACAGCTCGAACACGCGCGCCAGCTCCG 71120

RESULT 8
OSJN00061/c OSJN00061
133868 bp DNA linear
PIN 12-III-2002

DEFINITION Oryza sativa genomic DNA, chromosome 4, BAC clone: USUNBD0031E11, complete sequence.

ACCESSION AL606629

KEYWORDS	HTG.
SOURCE	Oryza sativa.
ORGANISM	Oryza sativa

Euarystocia; viridiplantae; streptophyta; emmryophyta; iiracneophyta; spermatophyta; magnoliophyta; liliopsida; poales; poaceae; enthartoidae; Oryzae; Oryza.

AUTHORS
Fu, G., Wang, S. Y., Ren, S. X., Lv, G., Lin, W., Gu, W. Q., Zhu, G. F.,
Tu, Y. F., Jia, J., Yin, H. F., Zhang, Y., Cai, Z., Chen, J., Kang, H.,
Chen, Y., Shao, C. Y., Sun, Y., Hu, O. P., Zhang, X. L., Zhang, W.

Wang, L.-J., Ding, C.-W., Sheng, H.-H., Gu, J.-L., Chen, S.-Y., Ni, L., Zhu, F.-H., Han, B., Feng, Q., Huang, Y.-C., Li, Y., Zhu, J.-J., Zhao, Q., Hu, X., Liu, Y.-L., Mu, J., Yu, Z., Chen, L., Fan, D.-L., Meng, Q.-J.,

Lǐ, C., Lǐ, T., Zhāng, Y., Hú, H., Jiā, P.-X., Qián, Y.-M., Yīng, K.,
Zhōu, B., Chén, Z.-H., Hào, P., Zhāng, L., Wú, M., Zhāng, R.-Q., Guān, J.-P.
and Hōng, G.-F.

TITLE Direct Submission
JOURNAL Submitted (27-JUN-2002) Han Bin, National Center for Gene Research,

intron 4300..4356
 /gene="T20K18.20"
 /number=1
 exon 4357..4452
 /gene="T20K18.20"
 /number=2
 intron 4453..4822
 /gene="T20K18.20"
 /number=2
 exon 4823..5035
 /gene="T20K18.20"
 /number=3
 gene 5620..8690
 /gene="T20K18.30"
 cds 5620..8690
 /gene="T20K18.30"
 /note="similarity to other hypothetical proteins -
 Arabidopsis thaliana"
 /codon_start=1
 /product="putative protein"
 /protein_id="CAB40985.1"
 /db_xref="GI:4586244"
 /translation="MEVKGPEKLMSPVSELPYFLLLLGLVTKALLIGPSSAIL
 VGNSCVITGLMPAHITWYICLARCIPKFEELCEFLHLLISTKYLFPVTSRTKIG
 LVKTLALVFLPPLLPVAGIVSGLEGLAYGFTPLMAFEGVBSVTSKCYCF
 VDLDFPCFHSYFSYMDLEKREWSADVELEIKLRPLSCILASLIGAVDVLITAV
 AVYKSPYMLIKGMRLELDVREGPFLSEVCVPAGLAILMPLAVAGAVASVLS
 FLIGYSGVIVHOEDSPRMGLNYIIAASVLPREYNDLLYREGTSLPCCVKTET
 VHGRILGESKRVADLSKSSSSLSGLVSEGRITLKAITLYKPIQVWMLKSEVN
 GILRLDGLIDVDEELVGNSSKLYIKLPAMVLOCLLASAKNSGLVITGVE
 LTELNSPRDKVFMVPLIMKEQIKMLKLTDEDFCLKLVWCKUEREDMDNTG
 PPSSTVRKAQLQAIIRLQGVAMSRIPTRRRFPMVLVKLYIEALEMGAGNRAG
 GLKXNSDOTGGLDRTETPDMGV"
 /number=1
 /gene="T20K18.30"
 /number=1
 intron 5714..6010
 /gene="T20K18.30"
 /number=1
 exon 6011..6403
 /gene="T20K18.30"
 /number=2
 intron 6404..6445
 /gene="T20K18.30"
 /number=2
 exon 6446..6536
 /gene="T20K18.30"
 /number=3
 intron 6537..6629
 /gene="T20K18.30"
 /number=3
 exon 6630..6918
 /gene="T20K18.30"
 /number=4
 intron 6919..7293
 /gene="T20K18.30"
 /number=4
 exon 7294..7400
 /gene="T20K18.30"
 /number=5
 intron 7401..7524
 /gene="T20K18.30"
 /number=5
 exon 7525..7690
 /gene="T20K18.30"
 /number=6
 intron 7691..7764
 /gene="T20K18.30"
 /number=6
 exon 7765..7960
 /gene="T20K18.30"
 /number=7

intron 7961..8058
 /gene="T20K18.30"
 /number=7
 exon 8059..8299
 /gene="T20K18.30"
 /number=8
 intron 8300..8494
 /gene="T20K18.30"
 /number=9
 exon 8495..8690
 /gene="T20K18.30"
 /number=9
 gene 11412..12269
 /gene="T20K18.40"
 cds 11412..12269
 /gene="T20K18.40"
 /note="similarity to hypothetical proteins Arabidopsis
 thaliana"
 /codon_start=1
 /product="putative protein"
 /protein_id="CAB40986.1"
 /db_xref="GI:4586245"
 /translation="MSLRSEGSTAEKITEDEPTVYKTAQSSVTCTIYQAHMVGFEVRV
 LMSKLMNHSTLVNVTGVQGMNYCCVDLPKHFMYKKYKSFSEVSENOVDVMDR
 SAKTNGGPEPSDDYVVALVSEEVLLGLDHKKAKPKTKSRPSLVDAALFYKENV
 GKLFSTAKKHDRREHIEVSESTAKEREMVIVDGLVAVRVNQLQMFKNQTV
 LVDEPVPVFPVDVDMVLFSTGTGHLGFIKPESEGETSNETNCSASSSSSEFC
 LFLAWKLE"
 /number=1
 /gene="T20K18.40"
 /number=1
 gene 13159..14844
 /gene="T20K18.50"
 complement (13159..14844)
 /gene="T20K18.50"
 complement (13159..14844)
 /gene="T20K18.50"
 /note="contains EST gb:R65054, T12953"
 /codon_start=1
 /product="putative protein"
 /protein_id="CAB40987.1"
 /db_xref="GI:4586246"
 /translation="WSGRSAFSRTGGRFPEVLGQANVLSIGSFVIVGVFTII
 AATPEDEPLFHPBQKITTLTNSNATLKDDSVKGEOPMANQAFAGFINIAD
 VETSENDSDGNQDQDINIPIQDQPEVFLMMATMEKFRDQSHFYKGRKVIYEGSS
 SSCDMARYRPKDGAAAFYDQYRFRVIEKSGNSVSMGIGEVHSGVNAKRRKPGF
 RNSGGKVDQFALPVGBAVNDLSLVSESEVPEGHVLYVSGGDRCKSNMFLMSF
 LCAGEAYILNRTLVMDTLCLSVYTLISQNEGKDRFRYPDEPHKEASMLDOVO
 FMADGKMYKKNGLKHLVEDFRVTPKLVVDVKTLLMKRTGYEPDNYRVCEGET
 ESVVORPNNLLMKSKRLMEIVSAISRLNMVDYDIIHERGOKARKNEVMPULKDTSP
 SSILSTQDKIEQGNLYIATNEPELSEFNFLLKDKYKXKLTDEFDFLWDESESEWVSET

Query Match 11.5% Score 70; DB 8; Length 79375;
 Best Local Similarity 56.5%; Pred. No. 3.5e-08;
 Matches 130; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 50 AATACTCAAGAACGGAATGTAACATCATGTTTGAATCTTGAATAATCAACCATGCC 109
 DB 68973 AAGAGCAGCAATCCACACACCTCATCTTGTGAACCTTAAATTAACATACC 69032
 QY 110 GTCTGAATTGGTGTGTAATCAATCATGCTCTGCAAGCTTTTCCACATTTTGCAC 169
 DB 69033 ATCTCATCAACATCACTTATCATCTTCTTGGCAATCTTCAAGTGTCTCCCTTG 69092
 QY 170 TCACATACCCAGCTTCTCCAGAGCATTCAGCTCCCAATCCCAACCCATCCCG 229
 DB 69093 CTTCAACCCCATGACGAGACCGACCTCAGCTTCTTCAATCTGAATGAACCATCAC 69152
 QY 230 ATTCAAGTCAACACTTAAAGCCTTCGCAAGGCTACTGTCCTTCTTCT 279
 DB 69153 ATTCGATCAATACTTAAGGCTTCTGTATGCTCCTTCTTCTCTT 69202

DB 88310 GAGCGCATGACGCTCTGTAACGCTGGGTGAACCTGTCACATGTCAGCACCCGCTGCC 88369

QY 374 ATCCAGTTCACAAACATGTTTCCCACTAAGGAGCTCCCAATTTCTTCAAGGCTGAATGAC 433

DB 88370 GTTGGCGTGAAGCGGCGCATGTGGCGGAGCTCTCTCCCTGTGACGGGAT----- 88423

QY 434 AGACCCGATTTCTTGAGGACCAACCAATTCAGCTCTCCAGACTPAACGACCATCTTCATT 493

DB 88424 ---CCGAGCCTCTCCAGGACATCCCGTCAAGCTCTCCCGGATATCCGGCCGCGCGCTC 88480

QY 494 CTTGTGAGCTTCTGAAATACGCTTGCAAGTGGCT 531

DB 88481 ACGGTGAAACAGCTCGAACCCGCGGAGCTCCGCT 88518

RESULT 12

LOCUS ATAC009853/c 90112 bp DNA linear PLN 24-JAN-2001

INITIATION Arabidopsis thaliana chromosome III BAC F2103 genomic sequence,

complete sequence.

AC009853

AC009853.4 GI:12408714

HTG.

Arabidopsis thaliana.

Arabidopsis thaliana.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosid II; Brassicales; Brassicaceae; Arabidopsids.

1 (bases 1 to 90112)

Lin.X, Kaul,S., Town,C.D., Benito,M.-I., Greasy,T.H., Haas,B.,

Romling,C.M., Koo,H., Fujii,C.Y., Utezaack,T.R., Barnstead,M.E.,

Boman,C.L., White,O., Niernan,W.C. and Frazer,C.M.

Arabidopsis thaliana chromosome III BAC F2103 genomic sequence

Unpublished

2 (bases 1 to 90112)

Lin.X. and Kaul,S.

Direct Submission

Submitted (03-SEP-1999) The Institute for Genomic Research, 9712

Medical Center Dr., Rockville, MD 20850, USA, xlin@tigr.org

3 (bases 1 to 90112)

Lin.X.

Direct Submission

Submitted (24-JAN-2001) The Institute for Genomic Research, 9712

Medical Center Dr., Rockville, MD 20850, USA

On Jan 24, 2001 this sequence version replaced gi:12280791.

Addresses all correspondence to:

Xiaoying Lin

The Institute for Genomic Research

9712 Medical Center Dr.

Rockville, MD 20850, USA

e-mail: xlin@tigr.org

BAC clone F2103 is from Arabidopsis chromosome III and is near the

molecular marker m1357.

The orientation of the sequence is from SP6 to T7 end of the BAC

clone.

Genes were identified by a combination of three methods: Gene

prediction programs including GAIL (available by anonymous ftp

from archur.epm.ornl.gov), GeneFinder (Phil Green, University of

Washington), GENSCAN (Chris Burge,

http://www.cse.cmu.edu/~chris/burges/GENSCAN.html), and NetPlantGene

(http://www.cbs.dtu.dk/netgene/cbnetgene.html), searches of the

complete sequence against a peptide database and the Arabidopsis

EST database at TIGR (http://www.tigr.org/cdb/atc.html).

Annotated genes are named to indicate the level of evidence for

their annotation. Genes with similarity to other proteins are named

after the database hits. Genes without significant peptide

similarity but with EST similarity are named as 'unknown' proteins.

Genes without protein or EST similarity, that are predicted to be

than two gene prediction programs over most of their length are

annotated as 'hypothetical' proteins. Genes encoding tRNAs are

predicted by tRNAscan-SE (Sean Eddy,

http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are

FEATURES

source

identified by repeatmasker (Arian Smit,
http://ftp.genome.washington.edu/RM/RepeatMasker.html). Regions of
genomic sequence that are not annotated as genes but have predicted
exons by GAIL are annotated as misc features.

Location/Qualifiers

1..90112

/organism="Arabidopsis thaliana"

/cultivar="Columbia"

/db_xref="taxon:3702"

/chromosome="III"

/map="m1357"

/clone="F2103"

/complement(1..>2130)

/gene="F2103.1"

/note="similar to translation initiation factor EIF-2B

beta subunit GB:090511 [Fugu rubripes]"

complement(join(1..84,165..290,370..472,610..717,

841..970,1475..1762,1845..1999,2078..2130))

/gene="F2103.1"

complement(join(1..84,165..290,370..472,610..717,

841..970,1475..1762,1845..1999,2078..2130))

/gene="F2103.1"

/codon_start=1

/product="putative translation initiation factor EIF-2B

beta subunit, 3' partial"

/protein_id="AAF02141.1"

/db_xref="GI:6041832"

/translation="MPDVSTVEEPVNLKRRKIGSGATATYELLRSVISHORP

HANQAAALIDNKAIGEPPMAGSAVVAAPARSTPPSLQTLGTEPSAVPTTSSGA

DSDDDDVNDKSLTGKIPGPAAGVVAAPARSTPPSLQTLGTEPSAVPTTSSGA

DEESTADSSLTRELKDVIGVNLQLEINGCHEQLEAIEH1H0EVLTLSSG

RVLEFELCAKREKSRFRVFEVGGAPRGVGHIAELVARGQTVTDSAVFAMISR

VNMVILGAHVAANGGVIGPVMNNAALAAQHAVFVLASGHLCPYPHPNPEVL

NELRSPSELDFGEF"

complement(1..2807..4488)

/gene="F2103.2"

complement(join(1..2807..3557,4072..4488))

/gene="F2103.2"

complement(join(1..2807..3557,4072..4488))

/gene="F2103.2"

/note="unknown protein"

/codon_start=1

/protein_id="AAF02142.1"

/db_xref="GI:6041833"

/translation="MDSCLNQTALQFLPSSRSGDGGGFFVPAKRIQYSSMV

VAAQGRCPGSSINAPLEPRSAQGRFLRSYLKMKROLFHYADELQOLDDRSAA

IARMSISSSDASLHRAELKEKCTKAVQDVMKMLFYKSETRVLPVLSGCI

VNGRLIWPSSKDWELSTYSCDTLEIKSHVAIVGLRVNSCVTDNMTTOIKHLR

KYVASILYGYFLKASLSRHOECSLSDHSGYLKSPFGCSFTTGQIISKQOLR

HYISDPDETLQRCAPRTEEARNLIEKQSLALFGEESDEIVTSSFLKRLVLEAV

ARQTLMDTELVDGAYKLKENGNADEQGGKSI"

4572..4590

/note="exon predicted by xgtrail, quality

marginal shadowexon"

5782..5863

/note="exon predicted by xgtrail, quality

marginal shadowexon"

6523..6654

/note="exon predicted by xgtrail, quality good shadowexon"

complement(6846..8714)

/gene="F2103.3"

/note="similar to glucan endo-1-3-beta-glucosidase

precursor GB:P52409 [Triticum aestivum]"

complement(join(6846..7111,7250..8542,8624..8714))

/gene="F2103.3"

complement(join(7074..7111,7250..8542,8624..8675))

/gene="F2103.3"

/codon_start=1

/product="putative glucan endo-1-3-beta-glucosidase"

/protein_id="AAF02143.1"

/db_xref="GI:6041834"

/translation="MSLLHLPLALSILISVGAKRSGPGQINYGQGNLPPSSDVN

LTKSLNAKRVKLYDANPKILALANGTIDIVSVVNPPELLVNISKSSLSLSDWIRSNIL

```

PRYPPTKIRYLTVGNELIPSELKSLVPMARKIQRSKSLGVKKVYKVTTLAVD
LOSSEPPSSGGEFREDISGLIMKPMLOFLKRTKSLFVDVYPPFAMQODPHNDLDAI
FESTVNTVDDPVEDNITSLNHLFQMDIDAFVAMRVRVYPIRIIVAVTGPMPNDYVGI
GANIYAAATYNNRVVKKLAADPPVGTAPAPKGLVPAFVALVYENKOTGPGERHGL
LHPNOVAVGIDLSGKTEKESLPAPENNIDYKXKICWCAVAGANTKODGLSLVSC
CGNNICDPIORGRCPOCKPDLTVLHASVYASSVYAPRKIGGTCFNGLATQTIKDSY
GRCEPSPVTL"
complement(10454..>13692)
/gene="F2103.4"
/note="similar to putative glucosyltransferase GB:AA023884
[Arabidopsis thaliana]"
complement(join(10454..11503,11632..11745,11846..12136,
12423..12733,12991..>13692))
/gene="F2103.4"
complement(join(10871..11503,11632..11745,11846..12136,
12423..12733,12991..13692))
/gene="F2103.4"
/note="unknown protein"
/codon_start=1
/protein_id="AA02144.1"
/db_xref="GI:6041835"
/translation="MSRSONEFOQMMNKORNDNHDVLAAGDAFLVETRTPTV
DDPKDRIPTRTQOLSLYLKPKLASSFLTNGSEFLVYIPRANRIANDNPSSVS
SARLYRLKGLVTVVLLCFELALFKGMHFTPPVASAEVAVYVAMLEIRASY
LAPLQSLTNVICVLFLQSDVRLVLGCFWIKLRKRVASMEYPTLVGSGVBLE
DYPWVVOIYPMCKEKEVYOOSIGAVCMLEPRRMVVOYLDSESDVQDLIAEVOK
WOORGVRIYPRHLIRITGYKAGLKAAMNCEYVDYEFVALPADQPPADFLKTPV
HFKNELALVOTRMAFVNKDEMLTRLQNLINSFPEVQOQNGVFNPGNGGQ
VVRKALPDGCGGLERTVEDMDIAYRAHCGKRFYINDVKLCGLSPSYEYKQO
YRHSQPMQFLRCFDILRSKVSAAKAMITLFLKRLILPFTSFIDVFCVILPT
MEFPEANLPSWVVCYIPGINSILNIIIPARSFPIYVILFENTMSVTKFGAMISLF
KEDSSYEVNTVKLGSRSEADLVAAVSGSLVSTVIOSSSGGTELSKGAAGA
GKTKRRLVYRTETALAFILAAVSRLSKQGHFYFLFQGITFVIVGDLIGEVS
"
13730..13843
/note="exon predicted by xgrail, quality
marginal, shadowexon"
complement(14067..14107)
/note="exon predicted by xgrail, quality good"
complement(14758..14844)
/note="exon predicted by xgrail, quality good"
15297..15363
/rpc_family="(TAA)n"
complement(<15938..>18038)
/gene="F2103.5"
/note="contains helix-loop-helix DNA binding motif"
complement(join(15938..15981,16078..16159,16251..16466,
16590..16661,16787..16855,16945..17010,17137..17289,
17370..>18038))
/gene="F2103.5"
complement(join(15938..15981,16078..16159,16251..16466,
16590..16661,16787..16855,16945..17010,17137..17289,
17370..18038))
/note="unknown protein"
/codon_start=1
/protein_id="AA02164.1"
/db_xref="GI:6041855"
/translation="MENELPMNAGVGHPPMTSPSSSANTKYNMETOPVDSLRN
LFWKSTEOSIFDSALSLVSSFTPSNSNVGCGENYIKELIGKLDNIDITGI
TASNGNSCYATPMSSPPGSMETKTTPPAELSGDGFERLAARSCFGSRFSNR
NSPPIINNEPITITNEKMPVSSPVFKPLASHVPAGESSGELSRKRTKSKNSSA
VS6KPIIEEKEDPKRCKKSENGDKTSDIPKYIHVARAGQATDSHAEVRA
REKISRMKLOLVPCNCKVAKALMLDEIINVYOSLOQVEFLSMKLSVNTLDF
NMALSKDIPSSNNLMHQVLCIDSSSETLIGDHNNKLOLNDISSNVINLE
TSBTRSFISHLPTLAHFTDISQYSTFSEDDLSTIHGPAQNRLOELNGSSNQVS
HMKAEI"
complement(18567..18642)
/note="exon predicted by xgrail, quality excellent"
complement(19174..19221)
/note="exon predicted by xgrail, quality good"
misc_feature
complement(20307..20333)

```

```

Query Match 10.5%; Score 64.2; DB 8; Length 90112;
Best Local Similarity 57.1%; Pred. No. 1.5e-06;
Matches 117; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

OY 74 CATCATGTTTGAATTCCTGTAATAACCAATGCGGTGAATGGTGTGAATPACA 133
DB 66361 CATCATGTTTGAATTCCTGTAATAACCAATGCGGTGAATGGTGTGAATPACA 133

OY 134 AATCATGCTCTGCAAGTCTTTTCCACTATTTTACCTACATACCAGTCTTCCAGCAC 193
DB 66301 TATCATCCTCTTGCAATCTCTAGGTTCTTCTGTTTGAAGCCCAAGAACTATAC 66242

OY 194 GTATTCAGCTCTCTACATATCCCAACCCATCCCATTCAGTCAAACTTTAAAC 253
DB 66241 AGACCTTACCTCTTCCACCGTATGAACCCGCGGTTGTATCAAAACGTTGAAAC 66182

OY 254 CTTCGCAAGCTACGCTCTTCTTCT 278
DB 66181 CTCTCTCATGCTCTTCTCTCTCTCT 66157

RESULT 13
AC123526/c
LOCUS
DEFINITION
Oryza sativa chromosome 11 clone OSJNBa005906, *** SEQUENCING IN
PROGRESS ***, 8 ordered pieces.
ACCESSION
AC123526
VERSION
AC123526.1 GI:21240712
KEYWORDS
HTG; HTGS PHASE2.
SOURCE
Oryza sativa.
ORGANISM
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 173645)
Linton,E.W., Tanyolac,B., Young,S., Kavchok,S., Keizer,G.,
Bronzino,A. and Messing,J.
Unpublished
2 (bases 1 to 173645)
Linton,E.W., Tanyolac,B., Young,S., Kavchok,S., Keizer,G.,
Bronzino,A. and Messing,J.
Direct Submission
Submitted (29-MAY-2002) The Plant Genome Initiative at Rutgers -
Makeman Institute, Rutgers, The State University of New Jersey, 190
Frelinghuysen Road, Piscataway, New Jersey 08854, USA
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 40577: contig of 40577 bp in length
40578 40677: gap of unknown length
40678 49266: contig of 8589 bp in length
49267 49366: gap of unknown length
49367 77857: contig of 28491 bp in length
77858 77957: gap of unknown length
77958 85343: contig of 7366 bp in length
85344 92753: contig of 7310 bp in length
92754 92853: gap of unknown length
92854 99880: contig of 6727 bp in length
99881 99981: gap of unknown length
99982 169296: contig of 69516 bp in length
169297 173645: contig of 4349 bp in length.
FEATURES
Location/Qualifiers
1..173645
/organism="Oryza sativa"

```

/db_xref="taxon:4530"
 /chromosome="11"
 /clone="OSJNB0059J06"
 BASE COUNT 48257 a 38169 c 38571 g 47946 t 702 others
 ORIGIN

Query Match 10.4% Score 63.2; DB 2; Length 173645;
 Best Local Similarity 53.2% Pred. No. 3.1e-06;
 Matches 134; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 74 CATCATGTTTGAATTTCTGAAATCAACATGCGCTGAATGTCGTCTAATACCA 133
 |||||
 Db 81824 CATCATGCTTTAACTGTCTAAATTTATCTTCATTCGCGCATCTCTGTGAAAACCTT 81765

QY 134 AATCATGCTCTGACGCTTTTTCACATATTTTACCTGACATATCCAGTCTTCCAGAC 193
 |||||
 Db 81764 GAGCATTTCTACACCTCTCTATGAGCGAGCGTGTGTGCAAAACCACTTCTCATAC 81705

QY 194 GTATTCAAGCTCTTCACATCCCGCCACACCATCCCATTCAGTCAAGTCAACCTTTAAAGC 253
 |||||
 Db 81704 ATTCTGACGCTTCAGAGGAGCTTATGATTCATTCGCGCTGGAACAGCTGAAAGC 81645

QY 254 CTTCGCAAGTCTACATGCTCTTCTTTCACCGCGCGGTAATGACCAATTCCTCTCTC 313
 |||||
 Db 81644 TTCTCTTAGTCTTCCCGAGCTCGCTTCTCTTCTTAATGACAGAGCTTCATCTTC 81585

QY 314 CTCTTCGTACCC 325
 |||||
 Db 81584 TACCTCGTTGCC 81573

RESULT 14
 LJA251808/c 631 bp mRNA linear PLAN 28-JAN-2002
 LOCUS LJA251808
 DEFINITION Locus japonicus mRNA for partial calcium-binding protein (cbp1 gene).

ACCESSION AJ251808.2 GI:18413494
 VERSION AJ251808.2
 KEYWORDS calcium-binding protein; cbp1 gene.
 SOURCE Locus japonicus.
 ORGANISM Locus japonicus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Locus.

REFERENCE 1
 Webb, K.J., Skot, L., Nicholson, M.N., Jorgensen, B. and Mizen, S.
 Mesorhizobium loti increases root-specific expression of a calcium-binding protein homologue identified by promoter tagging in Locus japonicus

JOURNAL MoL. Plant Microbe Interact. 13 (6), 606-616 (2000)
 MEDLINE 20289072
 PUBMED 10830260
 REFERENCE 2
 Skot, L.
 Direct Submision
 Submitted (10-DEC-1999) Skot L., Soils And Agroecology Department,
 Institute Of Grassland And Environmental Research, Plas Gogerddan,
 Aberystwyth, Ceredigion, SY23 3EB, UNITED KINGDOM
 Revised by [4]
 3 (bases 1 to 631)

REMARK 3
 Skot, L.
 Direct Submision
 Submitted (15-JAN-2002) Skot L., Soils And Agroecology Department,
 Institute Of Grassland And Environmental Research, Plas Gogerddan,
 Aberystwyth, Ceredigion, SY23 3EB, UNITED KINGDOM
 On Jan 29, 2002 this sequence version replaced g1:6580548.

COMMENT 1. 631
 Location/Qualifiers
 source
 /organism="Locus japonicus"
 /cultivar="Gifu B-129-S9"
 /db_xref="taxon:34305"
 /feature_type="root"

gene
 CDS

/country="Japan"
 1. 631
 /gene="cbp1"
 <1. 453
 /gene="cbp1"
 /codon_start=1
 /product="calcium-binding protein"
 /protein_id="CAB3264.2"
 /db_xref="GI:18413495"
 /db_xref="SPRMBL:O9SCA1"
 /translation="MDPTELKRVQMPDRNDGRITKLELNDLSLENIGIPIDKELTQ
 MIERIDVNGCCVDIDFEGELYSIMDERDEEDMRFAFVFDNGGQFIVEELRTV
 LASGIKQGRVDECKMKIMKVDVDSGMDYKFKFMKGGGSALT"
 451. 631
 3' UTR

BASE COUNT 178 a 119 c 173 g 161 t
 ORIGIN

Query Match 10.2% Score 62.2; DB 8; Length 631;
 Best Local Similarity 54.8% Pred. No. 2.4e-06;
 Matches 124; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 51 ATAATCAAGACGGGAATGTAATCATGTTTGAATTTCTGAAATCAACATGCGC 110
 |||||
 Db 446 AGAGCGCTAAGCCACCACTTATCTTCTGTAATCTCTTAATCAACATACCA 387

QY 111 TGTGAATGATGCTAATACCAATCATGCTCTGAGTCTTTTCCACTATTTACCT 170
 |||||
 Db 386 TCTCATCAATCATCACTTATGATCATCTTCTGCAATCTTCAGAGTCTCCCTGCT 327

QY 171 CACATPACCACTCTCCAGCAGCATATTCAGCTCTTCACATCCCGCCACCATCCCA 230
 |||||
 Db 326 TTGATCCCGAGTGAAGCAGAACCGCTCACTCTCCACGATGATGACCATCCCG 267

QY 231 TTCAAGTCAACACTTTAAAGCCTTGCAAGTCACTGCTTCTTC 277
 |||||
 Db 266 TTGTGATGAAGACGTGAAGAGCTCCTCATGCTCTCTCTCATC 220

RESULT 15
 AP004167/c 119048 bp DNA linear HTG 21-MAR-2002
 LOCUS AP004167
 DEFINITION Oryza sativa (japonica cultivar-group) chromosome 2 clone
 OJ1674.E08, *** SEQUENCING IN PROGRESS ***. In ordered pieces.
 ACCESSION AP004167.1 GI:15594179
 VERSION AP004167.1
 KEYWORDS HTG; HTGS PHASE2.
 SOURCE HTG: Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
 clone:OJ1674.E08.
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1
 Sasaki, T., Matsumoto, T. and Yamamoto, K.
 Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC
 clone:OJ1674.E08
 Published Only in Database (2001)

JOURNAL 2 (bases 1 to 119048)
 Sasaki, T., Matsumoto, T. and Yamamoto, K.
 Direct Submision
 Submitted (12-SEP-2001) Takuji Sasaki, National Institute of
 Agrobiological Sciences, Rice Genome Research Program, Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail:tsasaki@nias.affrc.go.jp, URL:http://snp.dna.affrc.go.jp/,
 Tel:81-298-38-7441, Fax:81-298-38-7468)

REMARK The nucleotide sequence of this BAC clone was generated by
 combining Monsanto and RGP-Japan sequencing data.
 NOTE: It currently consists of 1 contigs. Gaps between the contigs
 are represented as runs of N. The order of the pieces is believed
 to be correct as given, however the sizes of the gaps between them
 are based on estimates that have provided by the submitter. This
 sequence will be replaced by the finished sequence as soon as it is

sequence will be replaced by the finished sequence as soon as it is

available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES

source

1. 119048

/organism="Oryza sativa (japonica cultivar-group)"

/cultivar="Hippobare"

/db_xref="taxon:39947"

/chromosome="2"

/clone="OJ1674 E08"

BASE COUNT 31432 a 27160 c 27460 g 32946 t 50 others

ORIGIN

Query Match 9.9%; Score 60.2; DB 2; Length 119048;
 Best Local Similarity 54.8%; Pred. No. 2e-05;
 Matches 119; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

```

74 CATCATGTTTGAATTCTTGAATAATCAACCAATGCCGCTGAATTGTCGTATATACCA 133
   |||||
118187 CATCATGCTCTGAGACTGAGAGAGTGAAGCGCGCGCTGCCCTGTCAGCTGGCC 118128
   |||||
134 AATCATGCTCTGAGACTGCTTTTCCACTATTTCACCTCACATACCAAGTCTTCCAGCAC 193
   |||||
118127 GATCATGCGCGCGCGAGTCTCCGCTGTCCGCGCTCTGTGATCCCAAGCAAGCCAGCAC 118068
   |||||
194 GATTCAGCTCTCTCACATGCCCAACCCATCCCATTCAGTCAAAACACTTAAAGC 253
   |||||
118067 GGGGCCAGCTGTCCAGCGTATGAACCGGTGCGCGTCCGCTCGAAGACGTCGAACGC 118008
   |||||
254 CTTGCAAGTCACTGTCTTCTTTCACGCGCGCGC 290
   |||||
118007 CTCGCGCATGCTGCTCATCCAGCGCGCGC 117971

```

Search completed: June 29, 2003, 07:04:37
 Job time : 1250 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2003, 05:30:29 ; Search time 159 Seconds
(without alignments)
8625.578 Million cell updates/sec

Title: US-10-021-323-13

Perfect score: 609
Sequence: 1 ggaatgaatcaactttt.....tgaactncaactaag 609

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

1) number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*

1: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred.. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	81.2	13.3	381	ABQ85296
2	53.6	8.8	969	AAQ47473
3	53.6	8.8	972	AAQ47473
4	42.8	7.0	630	AAQ42498
5	41.6	6.8	5928	AAQ42498
6	41.6	6.8	5928	AAQ42498
7	40	6.6	916	AAQ42498
8	40	6.6	1052	AAQ42498
9	40	6.6	1052	AAQ42498

C 10	40	6.6	1052	21	AAQ88238
C 11	40	6.5	1052	24	AAQ84424
C 12	39.6	6.5	887	21	AAQ35926
C 13	39.2	6.4	1527	23	AAQ80811
C 14	38.8	6.4	688	21	AAQ52200
C 15	38.8	6.4	2876	22	AAQ17907
C 16	38.6	6.3	843	21	AAQ31085
C 17	38.4	6.3	369	24	AAQ85512
C 18	38.4	6.3	889	20	AAQ81537
C 19	38.4	6.3	889	21	AAQ8239
C 20	38.4	6.3	1035	21	AAQ8239
C 21	38.4	6.3	1155	23	AAQ76389
C 22	38.4	6.3	5785	20	AAQ81535
C 23	38.4	6.3	5785	21	AAQ88236
C 24	38	6.2	7231	24	AAQ54324
C 25	37.8	6.2	966	20	AAQ60314
C 26	37.8	6.2	1371	24	AAQ84741
C 27	37.2	6.1	2231	24	AAQ95536
C 28	37	6.1	318	21	AAQ10202
C 29	36.8	6.0	515	21	AAQ32863
C 30	36.8	6.0	1529	18	AAQ63437
C 31	36.4	6.0	462	8	AAQ70612
C 32	36.4	6.0	462	11	AAQ02033
C 33	36.2	5.9	1082	24	AAQ41274
C 34	36.2	5.9	1082	24	AAQ41275
C 35	36.2	5.9	1965	20	AAQ60621
C 36	36.2	5.9	8748	23	AAQ12492
C 37	36.2	5.9	33923	22	AAQ67071
C 38	36	5.9	3630	23	AAQ27020
C 39	35.8	5.9	486	21	AAQ49350
C 40	35.8	5.9	487	22	AAQ05366
C 41	35.8	5.9	487	22	AAQ30961
C 42	35.8	5.9	487	22	AAQ15480
C 43	35.8	5.9	487	22	AAQ15673
C 44	35.8	5.9	487	22	AAQ05709
C 45	35.8	5.9	515	22	AAQ25076

ALIGNMENTS

RESULT 1	ABQ85296	standard; DNA; 381 BP.
ID	ABQ85296	
XX	ABQ85296;	
AC	ABQ85296;	
DT	05-SEP-2002	(first entry)
XX		
DE	Arabidopsis thaliana expressed polynucleotide SEQ ID NO 166.	
XX		
KM	Plant; Arabidopsis; transgenic; fungicide; insecticide; de.	
XX		
OS	Arabidopsis thaliana.	
PN	US2002062014-A1.	
PD	23-MAY-2002.	
XX		
PF	26-JAN-2001; 2001US-0770791.	
XX		
PR	27-JAN-2000; 2000US-178480P.	
XX		
PA	(GORL/) GORLACH J.	
PA	(LANY/) AN Y.	
PA	(HANT/) HAMILTON C M.	
PA	(PRIC/) PRICE J L.	
PA	(RAIN/) RAINES T M.	
PA	(YUY/) YU Y.	
PA	(RAME/) RAMEKA J G.	
PA	(PAGE/) PAGE A.	
PA	(WATH/) MATTHEW A V.	
PA	(LEDF/) LEDFORD B L.	

Human pp32 nucleot
Human cDNA differe
Arabidopsis thalia
DNA encoding novel
Arabidopsis thalia
Human cDNA sequenc
Human colon cancer
Arabidopsis thalia
Genomic sequence o
Human variant pp32
DNA encoding novel
Genomic sequence o
Human placenta var
Chemically treated
DNA encoding a hum
Human cDNA differe
Mouse ischaemic co
Human secreted pro
Arabidopsis thalia
Repressed stearyl-
Sequence encoding
Sequence encoding
Oligonucleotide fo
Oligonucleotide seq
Polynucleotide seq
Drosophila melanog
Human immune/haema
Drosophila melanog
Arabidopsis thalia
Arabidopsis thalia
Human brain expres
Human bone marrow
Probe #5413 for ge
Probe #5559 used t
Human genome-deriv
Human brain expres

PA (WOES/) WOESSNER J P.
 PA (HAAS/) HAAS W D.
 PA (GARC/) GARCIA C A.
 PA (KRICK/) KRICKER M.
 PA (SLAT/) SLATER T.
 PA (DAVIS/) DAVIS K R.
 PA (ALLEN/) ALLEN K.
 PA (HOFF/) HOFFMAN N.
 PA (HURB/) HURBAN P.
 PI Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
 PI Ramezka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
 PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N,
 PI Hurban P;
 DR WPI; 2002-479265/51.
 XX
 PT New nucleic acid sequences of Arabidopsis thaliana and their encoded
 products are useful to produce transgenic plants, to screen for
 biologically active agents such as fungicides and insecticides and in
 genetic studies
 PS Claim 1; SEQ ID NO 166; 18bp + Sequence listing; English.
 XX
 CC The invention relates to a novel nucleic acid of Arabidopsis thaliana
 CC comprising a sequence capable of hybridising under stringency to one of
 CC the 999 sequences referred to but not defined in the specification.
 CC (AB085131-AB086129). The nucleic acid sequences are useful to identify
 CC homologous or related genes, to produce compositions that modulate
 CC expression or function of the encoded protein, to map functional regions
 CC of the protein, to study associated physiological pathways, to
 CC genetically manipulate cells and plants. The encoded products are useful
 CC to screen for biologically active agents such as fungicides or
 CC insecticides and to elucidate biochemical pathways.
 XX
 SQ Sequence 381 BP; 137 A; 101 C; 42 G; 101 T; 0 other;
 Query Match 13.3%; Score 81.2; DB 24; Length 381;
 Best Local Similarity 61.2%; Pred. No. 2e-14;
 Matches 131; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
 QY 68 ATGTACATCATGTTTGAATTTGAAATCAACATCCGCTCAATGGTGTGTA 127
 Db 167 ATGTAGATCATGTTCTTGAATTTCAAGTCAACAAACATCAAGATTCGTGATG 226
 QY 128 ATTCGAATCATGCTCGAGCTTTTCACATATTTTCACTACATACCCAGTCTCC 167
 Db 227 AACTCGAATCATTTCTCCCAATTCACAGCTTGTGCTCTCCCAACCTTAGCGTTTC 286
 QY 188 CAGCAGTATTCAGCTCTACATCCGCCCAACCCATCCCATTCAGTCAACACTTT 247
 Db 287 CACACATTCAGCAAGCTCTCCGCTGAATATACCATCTCCGTTCCATCAGAACATTT 346
 QY 248 AAAAGCTTGGCAAGTCACTGTCTTTCTTCA 261
 Db 347 AAAGCCCTCGCAATCGCTTCATCATATATACGA 380
 RESULT 2
 AAC47473/c
 ID AAC47473 standard; DNA; 969 BP.
 AC AAC47473;
 XX
 DT 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 53949.
 XX
 KM Hybridisation assay; genetic mapping; gene expression control;
 KM protein identification; signal transduction pathway;
 KM metabolic pathway; promoter; termination sequence; ss.
 XX
 OS Arabidopsis thaliana.

XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0125264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134421.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.


```

PR 01-JUL-1999; 99US-0141842;
PR 01-JUL-1999; 99US-0142154;
PR 02-JUL-1999; 99US-0142055;
PR 06-JUL-1999; 99US-0142390;
PR 08-JUL-1999; 99US-0142803;
PR 09-JUL-1999; 99US-0142920;
PR 12-JUL-1999; 99US-0142977;
PR 13-JUL-1999; 99US-0143542;
PR 14-JUL-1999; 99US-0143624;
PR 15-JUL-1999; 99US-0144005;
PR 16-JUL-1999; 99US-0144085;
PR 16-JUL-1999; 99US-0144326;
PR 19-JUL-1999; 99US-0144335;
PR 19-JUL-1999; 99US-0144334;
PR 19-JUL-1999; 99US-0144335;
PR 20-JUL-1999; 99US-0144352;
PR 20-JUL-1999; 99US-0144632;
PR 20-JUL-1999; 99US-0144884;
PR 21-JUL-1999; 99US-0144814;
PR 21-JUL-1999; 99US-0145086;
PR 22-JUL-1999; 99US-0145089;
PR 22-JUL-1999; 99US-0145192;
PR 23-JUL-1999; 99US-0145145;
PR 23-JUL-1999; 99US-0145218;
PR 23-JUL-1999; 99US-0145224;
PR 26-JUL-1999; 99US-0145276;
PR 27-JUL-1999; 99US-0145913;
PR 27-JUL-1999; 99US-0145918;
PR 28-JUL-1999; 99US-0145919;
PR 28-JUL-1999; 99US-0145951;
PR 02-AUG-1999; 99US-0146386;
PR 02-AUG-1999; 99US-0146388;
PR 02-AUG-1999; 99US-0146389;
PR 03-AUG-1999; 99US-0147038;
PR 04-AUG-1999; 99US-0147204;
PR 04-AUG-1999; 99US-0147302;
PR 05-AUG-1999; 99US-0147192;
PR 05-AUG-1999; 99US-0147260;
PR 06-AUG-1999; 99US-0147303;
PR 06-AUG-1999; 99US-0147416;
PR 09-AUG-1999; 99US-0147493;
PR 09-AUG-1999; 99US-0147935;
PR 10-AUG-1999; 99US-0148171;
PR 11-AUG-1999; 99US-0148319;
PR 12-AUG-1999; 99US-0148341;
PR 13-AUG-1999; 99US-0148565;
PR 13-AUG-1999; 99US-0148684;
PR 16-AUG-1999; 99US-0149368;
PR 17-AUG-1999; 99US-0149175;
PR 18-AUG-1999; 99US-0149426;
PR 20-AUG-1999; 99US-0149722;
PR 20-AUG-1999; 99US-0149723;
PR 20-AUG-1999; 99US-0149929;
PR 23-AUG-1999; 99US-0149902;
PR 23-AUG-1999; 99US-0149930;
PR 25-AUG-1999; 99US-0150566;
PR 26-AUG-1999; 99US-0150884;
PR 27-AUG-1999; 99US-0151065;
PR 27-AUG-1999; 99US-0151066;
PR 27-AUG-1999; 99US-0151080;
PR 30-AUG-1999; 99US-0151303;
PR 31-AUG-1999; 99US-0151438;
PR 01-SEP-1999; 99US-0151930;
PR 07-SEP-1999; 99US-0152363;
PR 10-SEP-1999; 99US-0153070;
PR 13-SEP-1999; 99US-0153758;
PR 15-SEP-1999; 99US-0154018;

```

```

PR 16-SEP-1999; 99US-0154039;
PR 20-SEP-1999; 99US-0154779;
PR 22-SEP-1999; 99US-0155139;
PR 23-SEP-1999; 99US-0155486;
PR 24-SEP-1999; 99US-0155659;
PR 28-SEP-1999; 99US-0156458;
PR 29-SEP-1999; 99US-0156596;
PR 04-OCT-1999; 99US-0157117;
PR 05-OCT-1999; 99US-0157753;
PR 06-OCT-1999; 99US-0157865;
PR 07-OCT-1999; 99US-0158029;
PR 08-OCT-1999; 99US-0158232;
PR 12-OCT-1999; 99US-0158369;
PR 13-OCT-1999; 99US-0158293;
PR 13-OCT-1999; 99US-0158294;
PR 13-OCT-1999; 99US-0158295;
PR 14-OCT-1999; 99US-0159329;
PR 14-OCT-1999; 99US-0159330;
PR 14-OCT-1999; 99US-0159331;
PR 14-OCT-1999; 99US-0159337;
PR 14-OCT-1999; 99US-0159637;
PR 14-OCT-1999; 99US-0159638;
PR 18-OCT-1999; 99US-0159584;
PR 21-OCT-1999; 99US-0160741;
PR 21-OCT-1999; 99US-0160767;
PR 21-OCT-1999; 99US-0160768;
PR 21-OCT-1999; 99US-0160770;
PR 21-OCT-1999; 99US-0160814;
PR 21-OCT-1999; 99US-0160815;
PR 22-OCT-1999; 99US-0160980;
PR 22-OCT-1999; 99US-0160981;
PR 22-OCT-1999; 99US-0160989;
PR 22-OCT-1999; 99US-0161004;
PR 25-OCT-1999; 99US-0161405;
PR 25-OCT-1999; 99US-0161406;
PR 26-OCT-1999; 99US-0161359;
PR 26-OCT-1999; 99US-0161360;
PR 26-OCT-1999; 99US-0161361;
PR 26-OCT-1999; 99US-0161361;
PR 28-OCT-1999; 99US-0161920;
PR 28-OCT-1999; 99US-0161921;
PR 28-OCT-1999; 99US-0161993;
PR 29-OCT-1999; 99US-0162142;

```

Query Match 8.8%; Score 53.6; DB 21; Length 969;
 Best Local Similarity 52.7%; Pred.No. 7.6e-06;
 Matches 116; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

```

QY 74 CATCATGTTTGAATTCCTGAATCAACCATGCGCTGAATTGCTGTAAATACCA 133
DB 708 CATCATGTAAGAAATCTCTTGTAAATGCTCTACCATCAATCTGCATCACTTGCA 649
QY 134 AATCATGCTCTGCAAGTCTTTTCCACTATTTCACCTCACATACCCAGCTTCCAGCAC 193
DB 648 AATCATCTTCTTACACACATCAAGGTCTCTCCCTTGCTGATGCCAAGAGACATCAC 589
QY 194 GTATTCAGCTCTCTCATCATCCCAACCCATCCCATTCATCAAGTCAACACTTTAAAC 253
DB 588 AGATTTCACTCTCCACAGATATTAACCATCTCCGCTTGTGTAACACGTTAAACGC 529
QY 254 CTTCGAAGTCACTGTCTTCTTTCACCGCGCGCGCA 293
DB 528 ATCTTCAATATCTCTCTCTCTGTTTGGCAGGTGTA 489

```

RESULT 3
 AAC34059/c
 ID AAC34059 standard; DNA; 972 BP.
 XX
 AC AAC34059;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 5304.
 XX

KM	Hybridization assay; genetic mapping; gene expression control;
KW	protein identification; signal transduction pathway;
XX	metabolic pathway; promoter; termination sequence; ss.
OS	Arabidopsis thaliana.
PN	EPI033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-0301439.
XX	
PR	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0125348.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.
PR	19-APR-1999; 99US-0130077.
PR	23-APR-1999; 99US-0130449.
PR	23-APR-1999; 99US-0130510.
PR	23-APR-1999; 99US-0130891.
PR	28-APR-1999; 99US-0131449.
PR	30-APR-1999; 99US-0132048.
PR	30-APR-1999; 99US-0133407.
PR	05-MAY-1999; 99US-0133484.
PR	06-MAY-1999; 99US-0133485.
PR	06-MAY-1999; 99US-0133486.
PR	07-MAY-1999; 99US-0133487.
PR	11-MAY-1999; 99US-0133863.
PR	14-MAY-1999; 99US-0134256.
PR	14-MAY-1999; 99US-0134218.
PR	14-MAY-1999; 99US-0134219.
PR	14-MAY-1999; 99US-0134221.
PR	18-MAY-1999; 99US-0134370.
PR	18-MAY-1999; 99US-0134768.
PR	19-MAY-1999; 99US-0134941.
PR	20-MAY-1999; 99US-0135124.
PR	21-MAY-1999; 99US-0135353.
PR	24-MAY-1999; 99US-0135629.
PR	25-MAY-1999; 99US-0136021.
PR	27-MAY-1999; 99US-0136392.
PR	28-MAY-1999; 99US-0136782.
PR	01-JUN-1999; 99US-0137222.
PR	03-JUN-1999; 99US-0137502.
PR	04-JUN-1999; 99US-0137502.
PR	07-JUN-1999; 99US-0137724.
PR	08-JUN-1999; 99US-0138094.
PR	10-JUN-1999; 99US-0138540.
PR	10-JUN-1999; 99US-0138847.
PR	14-JUN-1999; 99US-0139119.
PR	16-JUN-1999; 99US-0139452.
PR	17-JUN-1999; 99US-0139453.
PR	17-JUN-1999; 99US-0139492.
PR	18-JUN-1999; 99US-0139454.
PR	18-JUN-1999; 99US-0139455.
PR	18-JUN-1999; 99US-0139456.
PR	18-JUN-1999; 99US-0139457.
PR	18-JUN-1999; 99US-0139458.
PR	18-JUN-1999; 99US-0139459.
PR	18-JUN-1999; 99US-0139460.
PR	18-JUN-1999; 99US-0139461.
PR	18-JUN-1999; 99US-0139462.
PR	18-JUN-1999; 99US-0139463.
PR	18-JUN-1999; 99US-0139750.
PR	18-JUN-1999; 99US-0139763.
PR	21-JUN-1999; 99US-0139817.
PR	22-JUN-1999; 99US-0139899.
PR	23-JUN-1999; 99US-0140353.
PR	23-JUN-1999; 99US-0140354.
PR	24-JUN-1999; 99US-0140623.
PR	28-JUN-1999; 99US-0140895.
PR	29-JUN-1999; 99US-0140991.
PR	30-JUN-1999; 99US-0141287.
PR	01-JUL-1999; 99US-0141842.
PR	01-JUL-1999; 99US-0141854.
PR	02-JUL-1999; 99US-0142055.
PR	06-JUL-1999; 99US-0142390.
PR	08-JUL-1999; 99US-0142803.
PR	09-JUL-1999; 99US-0142920.
PR	12-JUL-1999; 99US-0142977.
PR	13-JUL-1999; 99US-0143542.
PR	14-JUL-1999; 99US-0143624.
PR	15-JUL-1999; 99US-0144005.
PR	16-JUL-1999; 99US-0144085.
PR	16-JUL-1999; 99US-0144085.
PR	19-JUL-1999; 99US-0144325.
PR	19-JUL-1999; 99US-0144331.
PR	19-JUL-1999; 99US-0144332.
PR	19-JUL-1999; 99US-0144333.
PR	19-JUL-1999; 99US-0144334.
PR	19-JUL-1999; 99US-0144335.
PR	20-JUL-1999; 99US-0144335.
PR	20-JUL-1999; 99US-0144632.
PR	20-JUL-1999; 99US-0144884.
PR	21-JUL-1999; 99US-0144814.
PR	21-JUL-1999; 99US-0145088.
PR	21-JUL-1999; 99US-0145088.
PR	22-JUL-1999; 99US-0145085.
PR	22-JUL-1999; 99US-0145087.
PR	22-JUL-1999; 99US-0145089.
PR	22-JUL-1999; 99US-0145119.
PR	23-JUL-1999; 99US-0145145.
PR	23-JUL-1999; 99US-0145218.
PR	23-JUL-1999; 99US-0145224.
PR	26-JUL-1999; 99US-0145276.
PR	27-JUL-1999; 99US-0145913.
PR	27-JUL-1999; 99US-0145918.
PR	27-JUL-1999; 99US-0145919.
PR	28-JUL-1999; 99US-0145951.
PR	28-JUL-1999; 99US-0145951.
PR	02-AUG-1999; 99US-0146386.
PR	

PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148664.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 22-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149950.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.

PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154739.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157765.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161004.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 7.0%; Score 42.8; DB 21; Length 630;
Best Local Similarity 47.7%; Pred. No. 0.012;
Matches 125; Conservative 0; Mismatches 137; Indels 0; Gaps 0;
QY 58 AAGACGGAGATGTATACATCATGTTTGAATTTGAAATCAACCAACCGCTGTGAT 117
DB 613 AAAAGGACCTTTTTCATCATTTGTCTAACTCATGTATTAACCTACATACCAT 554
QY 118 TGGTGTCTAATACCAATCATGCTCTGACGCTTTTTCACATATTTTCACTCATAC 177
DB 553 CAACATCACTTGATATCATCTTCTCTCATCTTCCCAAGCTTACCTTCTGAGTC 494
QY 178 CCAAGTCTCCACGACGATTTCAAGTCTCTCATCTCCCAACCCATCCCATTCAGT 237
DB 493 CCAAGGAGATTAACCGCTTCAATTCATCCACGTTATTAACCATCTCCGTTCCGAT 434
QY 238 CAACACTTAAGGCTTCCGAGGTCACTGCTTCTTCCACCGCGCGGTAAATGA 297
DB 433 CAATACGTTGAACGTTCTCTCATATCTTCTTCTCTTCTCTCTCTCTCTCTCTCA 374
QY 298 CCAATTCCT 319


```

Db      3793 CCTCMAACCCCATTTTCAACCCCTATCCCTCAACCTTATCTTCAAAACCTTCAT 3734
Qy      228 CCATTCAAGTCAAAAGCTTAAAGCCTTCGCAAGGTCACTGTCTTCTTTCACGCGG 287
Db      3733 CTTCCAAACCCCTACTCTCTCCAAACCTCATCTCCAAACCTCATCTCTCAAAACCCCA 3674
Qy      288 CGGTAAATGACCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 335
Db      3673 TCTTCAACCCCTATCTCTCTCAAAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3626

```

RESULT 7
AAV71743/C
ID AAV71743 standard; cDNA; 916 BP.

AC AAV71743;
XX
DT 15-MAR-1999 (first entry)

Human V3 loop HIV receptor P30/PHAPI cDNA.

KM HIV receptor; V3 loop; human immunodeficiency virus; retrovirus;
KM P30 protein; PHAPI; infection; therapy; diagnosis; ss.

OS Homo sapiens.

Key Location/Qualifiers
FT CDS 104..853
FT /*tag= a

WO9840480-A1.

17-SEP-1998.

12-MAR-1998; 98WO-EP01409.

12-MAR-1997; 97US-0040969.

(CNS) CEMT NAT RECH SCT.

(INSP) INST PASTEUR.

Briland J, Callebaut C, Guichard G, Hovanessian A;
PI Jacotot E, Krust B, Muller S;

WPI: 1999-034588/03.

P-PSDB; AAW84053.

New isolated V3 loop HIV receptor - comprises P95/nucleolin,
P40/PHAPI and P30/PHAPI proteins, used to develop products for the
treatment and prevention of HIV infection

Claim 17; Fig 49(12); 267bp; English.

This cDNA sequence codes for the P30 (or PHAPI) protein of the
newly identified V3 loop HIV receptor. This novel protein complex
receptor for HIV retroviruses consists of an association of 3

proteins named P95/nucleolin, P40/PHAPI and P30/PHAPI (see
AAW84052-54). These proteins were isolated from human CD4+ CEM

T-cell extracts using an affinity matrix containing either the
pseudopeptide 5 (KpsICH2)NPR)-template assembled synthetic peptide
or a synthetic V3 loop peptide (see AAW84055). P30 was identified

as PHAPI by amino acid sequence analysis. The invention also
concerns peptidic or non-peptidic molecules having the ability to
alter and/or prevent the binding of the novel HIV receptor to the

HIV retrovirus, and to pharmaceutical and diagnostic compositions
containing such molecules. Methods are provided for screening for
new active molecules, and to methods of screening genetic defects

in the expression of the V3 loop HIV receptor in individuals that
survive long-term HIV infection or who are HIV-resistant. Such
genetically defective polynucleotides can be used in gene therapy.

Sequence 916 BP; 296 A; 167 C; 275 G; 178 T; 0 other;

Query Match 6.6%; Score 40; DB 20; Length 916;
Best Local Similarity 46.7%; Pred. No. 0.1;
Matches 127; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

```

Qy      59 AACAAGGAAATGATTAATCATGTTTGAATTTGAAATTTGAAATACCAATGCGTGAAT 118
Db      886 ACAATAGGAATTTTCAAAATAGTTAATTCACCTTAGCATCATCATCTTCTCCCTCATCTTC 827
Qy      119 GGTGTGTAATACCAATATCATGCTTGTGAGTCTTTTCCACATTTTACCTCAATAC 178
Db      826 AGGTTCTGTTTTGCTTCTCTGACCCCTTCTTCTTACCAACCTCTTCTTCAATCTTCTC 767
Qy      179 CAGTCTCCAGACAGTATTCAGTCCCTACATCCCCCAACCATCCCATTCAGTAC 238
Db      766 GTCATCTACTCTTCATGTTTAACTTTTCACTTCTTCACTCTCTCTCTCTCTCTCTCT 707
Qy      239 AACAATTTAAAGCCTTCGCAAGGTCACTGTCTTCTTTCACCGCGCGTGAATGAC 298
Db      706 CTCTCTTCAACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 647
Qy      299 CAATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 330
Db      646 TTCAATCATACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 615

```

RESULT 8
AAT27712/C
ID AAT27712 standard; cDNA; 1052 BP.

AAT27712;

30-JUL-1996 (first entry)

Human pp32 cDNA.

pp32; cancer; diagnosis; therapy; antisense; cell proliferation;
KM lymphoid tumour; epithelial tumour; colon carcinoma;

KM prostate carcinoma; non-Hodgkin lymphoma; de.

OS Homo sapiens.

Key Location/Qualifiers

FT CDS 97..846

P-PSDB; AAR95900.

WO9610092-A1.

04-APR-1996.

28-SEP-1995; 95WO-US12414.

28-SEP-1994; 94US-0314503.

(UYGO) UNIV JOHNS HOPKINS.

Kuhajda FP, Pasternack GR;

WPI: 1996-200930/20.

P-PSDB; AAR95900.

New method of diagnosing cancer using pp32 cDNA - by detecting the
level of mRNA hybridizing to pp32 cDNA, also for inhibiting cell
proliferation and screening anti-cancer drugs.

Claim 1; Fig 10a; 129bp; English.

Human cDNA (AAT27712) cloned from HL-60 cells codes for a 32 kDa
protein, pp32 (AAR95900). The level of expression of pp32 correlates
with the malignant potential of lymphoid and epithelial tumours. The
cDNA clone was identified using a partial murine pp32 clone. The
human sequence may be used for specific diagnostic assays of tumour
tissue or in the creation of antisense expression vectors to inhibit

CC inducer of pp32 function comprising measuring protein phosphatase
 CC activity in cells cultured in the presence and absence of the compound.
 CC The methods are useful for treating cancer and for identifying agents
 CC which may be used to treat cancer. Human pp32 is a phosphoprotein which
 CC has been mapped to chromosome 15q22.3-q23. The present sequence
 CC represents the human pp32r1 nucleotide sequence from an example of
 CC the present invention.

XX Sequence 1052 BP; 326 A; 196 C; 316 G; 214 T; 0 other;
 SQ

Query Match 6.6%; Score 40; DB 21; Length 1052;
 Best Local Similarity 46.7%; Pred. No. 0.11;
 Matches 127; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 59 AGAAGCGAATGTAATCATATGTTTGAATTTGAAATTAACAACATGCGTGAATT 118
 DB 879 ACAATAGGAATTTTAAATAAGTATATCCACTAGTATCATCTTCTCCATCTTC 820
 OV 119 GGTGTGTAAATACCAATATGCTCTGACGCTTTTCCACTTTTACCTACATACC 178
 819 AGGTTCCTGTTTGGCTTGGACCCCTTCTCTTCAACCAAGCTCTTCAATCTTCTC 760
 QY 179 CAGTCTCCAGACGATATCAAGCTCCACATCCGCCAACCACATCCCATCAAGTC 238
 DB 759 GTCATCTACCTCTCCATCGTTATTAACCTTTTCACTCTCTCTCTCCATCAAGTC 700
 QY 239 AAACACTTTAAAGCTTCCGCAAGTCACTGCTCTTCTTTCACCGCCGCGTAATGAC 298
 DB 699 CTCCTCTTACCTTCCCTCCCTCCATCTGCTCTCTCTCTTCCACTACGAGCATC 640
 QY 299 CAATTCCT 330
 DB 639 TTCATCATCT 608

RESULT 11
 ABR84424/c
 ID ABR84424 standard; cDNA; 1052 BP.
 AC ABR84424;
 XX
 DT 14-AUG-2002 (first entry)
 XX
 DE Human cDNA differentially expressed in granulocytic cells #995.
 XX
 KM Human; ss; granulocytic cell; DNA chip; bacterial infection;
 KM viral infection; parasitic infection; protozoal infection;
 KM fungal infection; sterile inflammatory disease; psoriasis;
 KM rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
 KM cardiac reperfusion injury; renal reperfusion injury; ARDS;
 KM adult respiratory distress syndrome; inflammatory bowel disease;
 KM Crohn's disease; ulcerative colitis; periodontal disease;
 KM granulocyte activation; chronic inflammation; allergy.
 XX
 XX Homo sapiens.
 OS
 XX
 PN MO200228999-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 03-OCT-2001; 2001MO-US30821.
 XX
 PR 03-OCT-2000; 2000US-237189P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
 XX
 DR WPI; 2002-435328/46.
 XX
 PT Detecting granulocyte activation by detecting differential expression
 PT of genes associated with granulocyte activation, which serves as
 PT diagnostic markers that is useful for monitoring disease states and

PT drug toxicity -
 XX
 XX Claim 1; SEQ ID No 995; 114bp; English.
 XX
 CC The invention relates to detecting (M1) granulocyte (GC) activation
 CC (GCA), by detecting the level of expression of gene(s) (Gg) identified by
 CC DNA chip analysis as given in the specification, and comparing
 CC the expression level to an expression level in an unactivated
 CC GC, where differential expression of Gg is indicative of GCA.
 CC Also included are modulating (M2) GA by contacting GC with an agent
 CC that alters the expression of at least one gene in Gg; (2) screening (M3)
 CC for an agent capable of modulating GCA or an inflammation (especially
 CC chronic) in a tissue, an allergic response in a subject, exposure of a
 CC subject to a pathogen or sterile inflammatory disease, by detecting the
 CC level of expression in a sample of the tissue of gene(s) from Gg, where
 CC the level of expression of the gene is indicative of inflammation;
 CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,
 CC or allergic response in a subject, exposure of a subject to a pathogen
 CC or sterile inflammatory disease, by contacting a tissue having
 CC inflammation with an agent that modulates the expression of gene(s)
 CC from Gg in the tissue. M1 is useful for detecting GCA; M2 is useful for
 CC modulating GA; M3 is useful for screening an agent capable of modulating
 CC GCA preferably in an inflammation in a tissue; M4 is useful for
 CC detecting an inflammation (especially chronic) in a tissue, an allergic
 CC response in a subject, exposure of a subject to a pathogen or sterile
 CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
 CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
 CC reperfusion injury, ARDS, adult respiratory distress syndrome,
 CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
 CC periodontal disease, also bacterial infection, viral infection, and
 CC parasitic infection, protozoal infection, fungal infection and M5 is
 CC useful for treating one of the above conditions. The present
 CC sequence represents a gene differentially expressed in granulocytes.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 1052 BP; 326 A; 196 C; 316 G; 214 T; 0 other;
 Query Match 6.6%; Score 40; DB 24; Length 1052;
 Best Local Similarity 46.7%; Pred. No. 0.11;
 Matches 127; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 59 AGAAGCGAATGTAATCATATGTTTGAATTTGAAATTAACAACATGCGTGAATT 118
 DB 879 ACAATAGGAATTTTAAATAAGTATATCCACTAGTATCATCTTCTCCATCTTC 820
 QY 119 GGTGTGTAAATACCAATATGCTCTGACGCTTTTCCACTTTTACCTACATACC 178
 DB 819 AGGTTCCTGTTTGGCTTGGACCCCTTCTCTTCAACCAAGCTCTTCAATCTTCTC 760
 QY 179 CAGTCTCCAGACGATATCAAGCTCCACATCCGCCAACCACATCCCATCAAGTC 238
 DB 759 GTCATCTACCTCTCCATCGTTATTAACCTTTTCACTCTCTCTCTCCATCAAGTC 700
 QY 239 AAACACTTTAAAGCTTCCGCAAGTCACTGCTCTTCTTTCACCGCCGCGTAATGAC 298
 DB 699 CTCCTCTTACCTTCCCTCCCTCCATCTGCTCTCTCTCTTCCACTACGAGCATC 640
 QY 299 CAATTCCT 330
 DB 639 TTCATCATCT 608

RESULT 12
 AAC35926/c
 ID AAC35926 standard; DNA; 887 BP.
 AC AAC35926;

XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 11914.
DE Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 01-APR-1999; 99US-0126785.
PR 06-APR-1999; 99US-0127462.
PR 08-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0129845.
PR 21-APR-1999; 99US-0130047.
PR 23-APR-1999; 99US-0130449.
PR 28-APR-1999; 99US-0130510.
PR 30-APR-1999; 99US-0130891.
PR 04-MAY-1999; 99US-0131449.
PR 05-MAY-1999; 99US-0132407.
PR 06-MAY-1999; 99US-0132484.
PR 07-MAY-1999; 99US-0132485.
PR 08-MAY-1999; 99US-0132486.
PR 09-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139819.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140921.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145224.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.

ID	AAAC52200	standard; DNA; 688 BP.	PR	18-JUN-1999;	99US-0139461
XX	AAAC52200;		PR	18-JUN-1999;	99US-0139462
XX	18-OCT-2000	(first entry)	PR	18-JUN-1999;	99US-0139463
DT	18-OCT-2000		PR	18-JUN-1999;	99US-0139750
XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 70578.		PR	21-JUN-1999;	99US-0139763
XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 70578.		PR	22-JUN-1999;	99US-0139817
XX	Hybridisation assay; genetic mapping; gene expression control;		PR	23-JUN-1999;	99US-0140353
KW	protein identification; signal transduction pathway;		PR	23-JUN-1999;	99US-0140354
KW	metabolic pathway; promoter; termination sequence; ss.		PR	24-JUN-1999;	99US-0140695
XX			PR	28-JUN-1999;	99US-0140823
OS	Arabidopsis thaliana.		PR	29-JUN-1999;	99US-0140991
XX	EP1033405-A2.		PR	30-JUN-1999;	99US-0141287
PN	06-SEP-2000.		PR	01-JUL-1999;	99US-0141842
XX			PR	01-JUL-1999;	99US-0142154
XX			PR	02-JUL-1999;	99US-0142055
PD	25-FEB-2000; 2000EP-0301439.		PR	06-JUL-1999;	99US-0142390
XX			PR	08-JUL-1999;	99US-0142803
XX			PR	09-JUL-1999;	99US-0142820
PF			PR	12-JUL-1999;	99US-0142977
XX			PR	13-JUL-1999;	99US-0143324
XX			PR	14-JUL-1999;	99US-0143624
XX			PR	15-JUL-1999;	99US-0144005
XX			PR	15-JUL-1999;	99US-0144085
XX			PR	16-JUL-1999;	99US-0144086
XX			PR	16-JUL-1999;	99US-0144331
XX			PR	19-JUL-1999;	99US-0144331
XX			PR	19-JUL-1999;	99US-0144332
XX			PR	19-JUL-1999;	99US-0144333
XX			PR	19-JUL-1999;	99US-0144334
XX			PR	19-JUL-1999;	99US-0144334
XX			PR	19-JUL-1999;	99US-0144335
XX			PR	20-JUL-1999;	99US-0144632
XX			PR	20-JUL-1999;	99US-0144684
XX			PR	21-JUL-1999;	99US-0144818
XX			PR	21-JUL-1999;	99US-0145086
XX			PR	21-JUL-1999;	99US-0145088
XX			PR	22-JUL-1999;	99US-0145087
XX			PR	22-JUL-1999;	99US-0145089
XX			PR	22-JUL-1999;	99US-0145145
XX			PR	23-JUL-1999;	99US-0145145
XX			PR	23-JUL-1999;	99US-0145218
XX			PR	23-JUL-1999;	99US-0145224
XX			PR	26-JUL-1999;	99US-0145276
XX			PR	27-JUL-1999;	99US-0145513
XX			PR	27-JUL-1999;	99US-0145518
XX			PR	27-JUL-1999;	99US-0145519
XX			PR	28-JUL-1999;	99US-0145551
XX			PR	02-AUG-1999;	99US-0146388
XX			PR	02-AUG-1999;	99US-0146388
XX			PR	02-AUG-1999;	99US-0146388
XX			PR	03-AUG-1999;	99US-0147038
XX			PR	04-AUG-1999;	99US-0147204
XX			PR	05-AUG-1999;	99US-0147302
XX			PR	05-AUG-1999;	99US-0147322
XX			PR	05-AUG-1999;	99US-0147360
XX			PR	06-AUG-1999;	99US-0147303
XX			PR	06-AUG-1999;	99US-0147416
XX			PR	09-AUG-1999;	99US-0147493
XX			PR	09-AUG-1999;	99US-0147935
XX			PR	10-AUG-1999;	99US-0148171
XX			PR	11-AUG-1999;	99US-0148319
XX			PR	12-AUG-1999;	99US-0148341
XX			PR	13-AUG-1999;	99US-0148365
XX			PR	13-AUG-1999;	99US-0148864
XX			PR	16-AUG-1999;	99US-0149368
XX			PR	17-AUG-1999;	99US-0149175
XX			PR	18-AUG-1999;	99US-0149426
XX			PR	20-AUG-1999;	99US-0149723
XX			PR	20-AUG-1999;	99US-0149929
XX			PR	23-AUG-1999;	99US-0149902

PR	18-JUN-1999	99US-0139461
PR	18-JUN-1999	99US-0139462
PR	18-JUN-1999	99US-0139463
PR	18-JUN-1999	99US-0139750
PR	18-JUN-1999	99US-0139763
PR	21-JUN-1999	99US-0139817
PR	22-JUN-1999	99US-0139819
PR	23-JUN-1999	99US-0140353
PR	23-JUN-1999	99US-0140354
PR	24-JUN-1999	99US-0140695
PR	28-JUN-1999	99US-0140823
PR	29-JUN-1999	99US-0140991
PR	30-JUN-1999	99US-0142887
PR	01-JUL-1999	99US-0142814
PR	01-JUL-1999	99US-0142855
PR	02-JUL-1999	99US-0142905
PR	06-JUL-1999	99US-0142930
PR	09-JUL-1999	99US-0142820
PR	12-JUL-1999	99US-0142977
PR	14-JUL-1999	99US-0143542
PR	14-JUL-1999	99US-0143524
PR	15-JUL-1999	99US-0144005
PR	16-JUL-1999	99US-0144085
PR	19-JUL-1999	99US-0144086
PR	19-JUL-1999	99US-0144325
PR	20-JUL-1999	99US-0144331
PR	20-JUL-1999	99US-0144332
PR	21-JUL-1999	99US-0144384
PR	21-JUL-1999	99US-0144814
PR	21-JUL-1999	99US-0145086
PR	21-JUL-1999	99US-0145088
PR	22-JUL-1999	99US-0145087
PR	22-JUL-1999	99US-0145089
PR	22-JUL-1999	99US-0145089
PR	23-JUL-1999	99US-0145142
PR	23-JUL-1999	99US-0145145
PR	23-JUL-1999	99US-0145224
PR	26-JUL-1999	99US-0145576
PR	27-JUL-1999	99US-0145513
PR	27-JUL-1999	99US-0145518
PR	28-JUL-1999	99US-0145519
PR	02-AUG-1999	99US-0146386
PR	02-AUG-1999	99US-0146388
PR	02-AUG-1999	99US-0146389
PR	03-AUG-1999	99US-0146703
PR	04-AUG-1999	99US-0147204
PR	04-AUG-1999	99US-0147302
PR	05-AUG-1999	99US-0147192
PR	05-AUG-1999	99US-0147260
PR	06-AUG-1999	99US-0147716
PR	09-AUG-1999	99US-0147935
PR	09-AUG-1999	99US-0147935
PR	10-AUG-1999	99US-0148171
PR	11-AUG-1999	99US-0148319
PR	12-AUG-1999	99US-0148341
PR	13-AUG-1999	99US-0148565
PR	13-AUG-1999	99US-0148684
PR	16-AUG-1999	99US-0149168
PR	17-AUG-1999	99US-0149175
PR	18-AUG-1999	99US-0149426
PR	20-AUG-1999	99US-0149722
PR	20-AUG-1999	99US-0149723
PR	22-AUG-1999	99US-0149929
PR	22-AUG-1999	99US-0149902

```

PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161932.
PR 28-OCT-1999; 99US-0161933.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 6.4%; Score 38.8; DB 21; Length 688;
Best Local Similarity 70.3%; Pred. No. 0.2;
Matches 52; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 253 CCTTGCAGGTCATGCTCTTCTTTCACCGCGCGCGTGAATGACCAATTCTCTCTCT 312
DB 426 CCTCCGCGATCTCATGCTGCTTCTTTCACCAACATCATCATCTGCTTTCATCTCTCT 367
QY 313 CCTCTGCTCACA 326
DB 366 CCTCATCTTCAGCA 353

```

```

RESULT 15
AAH17907/c
ID AAH17907 standard; cDNA, 2876 BP.
XX
AC AAH17907;

```

```

XX 26-JUN-2001 (first entry)
DT Human cDNA sequence SEQ ID NO:17652.
DE Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX EPI074617-A2.
XX 07-FEB-2001.
PD 28-JUL-2000; 2000EP-0116126.
PF 29-JUL-1999; 99UP-0248036.
PR 27-AUG-1999; 99UP-0300253.
PR 11-JAN-2000; 2000UP-0118776.
PR 02-MAY-2000; 2000UP-0183767.
PR 09-JUN-2000; 2000UP-0241899.
XX
XX (HELI-) HELIX RES INST.
PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
PI WPI, 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
XX Claim 8; SEQ ID 17652; 2537pp + CD ROM; English.
PS
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination
CC of the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialized methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 2876 BP; 766 A; 709 C; 636 G; 705 T; 0 other;

```

```

Query Match 6.4%; Score 38.8; DB 22; Length 2876;
Best Local Similarity 48.2%; Pred. No. 0.4;
Matches 109; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 244 CTTTAAAGCTTTCGACAGTCTCTTCTTTCACCGCGCGCGTGAATGACCAATT 303
DB 2253 CTTAAATCATCGGTGTCATCATCTTTCATCATCATCTTCTCTCTCTCTCT 2194
QY 304 CCTCTCTCTCTCTTTCGACCAAGTGGTTCGAGATGATTAATAAGA 363
DB 2193 CTTGCTCTCTCTTTCATCTCTCTCTGCGAAGCGCTGCGCCATTATACAGCC 2134
QY 364 ACAAGATTCATCAAGTTCAAAACATGTTTCCACTAAGGCTCAATTCTTCAGGC 423

```

Db 2133 TGCATCTGTCTCCATGTTCAAGGCAATCCATGCTGCTTGAATTTCAATGACACGAAAGAT 2074
Qy 424 TGAATTGACAGACCCGATTCTCTGAGCAACCAATTCACTCTC 469
Db 2073 AGCAGTCTTTGGTGTCTGTACACAGGTCAAAGATGTTCCGTTTCA 2028

Search completed: June 29, 2003, 06:43:32
Job time : 161 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2003, 06:40:58 ; Search time 44 Seconds
(without alignments)
4244.685 Million cell updates/sec

Title: US-10-021-323-13

Perfect score: 609
Sequence: 1 ggaatgaatcaacttc.....tgaagcctacaataaag 609

Scoring table: IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 441362 seqs, 1513381 residues

1 number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCFUS.COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	51.6	8.5	7218 1	US-08-232-463-14 Sequence 14, Appl
2	40	6.6	1052 1	US-08-466-603-1 Sequence 1, Appl
3	40	6.6	1052 1	US-08-314-503A-1 Sequence 1, Appl
4	40	6.6	1052 1	US-08-468-066-1 Sequence 1, Appl
5	40	6.6	1052 2	US-08-466-717-1 Sequence 1, Appl
6	40	6.6	1052 2	US-08-466-717-1 Sequence 1, Appl
7	40	6.6	1052 5	PCF-US95-12414-1 Sequence 1, Appl
8	39.8	6.5	289 4	US-09-007-005-17 Sequence 17, Appl
9	39.8	6.5	289 4	US-09-244-786-17 Sequence 17, Appl
10	37.8	6.2	966 2	US-08-766-738-2 Sequence 2, Appl
11	37.8	6.2	966 4	US-09-262-610-2 Sequence 2, Appl
12	36.2	5.9	1965 4	US-09-178-252-26 Sequence 26, Appl
13	35.6	5.8	3211 2	US-08-574-959A-8 Sequence 8, Appl
14	35.6	5.8	3211 4	US-09-357-014-8 Sequence 8, Appl
15	35.6	5.8	3901 2	US-08-574-959A-6 Sequence 6, Appl
16	35.6	5.8	3901 4	US-09-357-014-6 Sequence 6, Appl
17	35.4	5.8	3489 2	US-08-728-323A-1 Sequence 1, Appl
18	35.4	5.8	3489 4	US-09-298-568-1 Sequence 1, Appl
19	35.4	5.8	33207 2	US-08-770-379-20 Sequence 20, Appl
20	35.4	5.8	33207 4	US-08-757-668A-20 Sequence 20, Appl
21	35.4	5.8	32207 4	US-09-230-371A-20 Sequence 20, Appl
22	34.4	5.6	856 4	US-09-171-517B-15 Sequence 15, Appl
23	34.4	5.6	4161 4	US-09-185-244-8 Sequence 8, Appl
24	34.4	5.6	4161 4	US-09-471-913-1 Sequence 1, Appl
25	34.4	5.6	13737 4	US-09-538-414-10 Sequence 10, Appl
26	33.6	5.5	2082 4	US-09-440-325A-2 Sequence 2, Appl
27	33.4	5.5	2277 1	US-08-676-967-2 Sequence 2, Appl

28	33.4	5.5	2277 1	US-08-676-974-2 Sequence 2, Appl
29	33.4	5.5	2277 2	US-09-098-487-2 Sequence 2, Appl
30	33.2	5.5	3624 1	US-07-951-715A-6 Sequence 6, Appl
31	33.2	5.5	3624 2	US-08-459-448A-6 Sequence 6, Appl
32	33.2	5.5	3624 3	US-08-459-595A-6 Sequence 6, Appl
33	33.2	5.5	3624 3	US-08-459-504B-6 Sequence 6, Appl
34	33.2	5.5	3624 3	US-08-459-444-6 Sequence 6, Appl
35	33.2	5.5	3624 3	US-09-053-549-7 Sequence 7, Appl
36	33.2	5.5	3624 4	US-09-547-422-6 Sequence 6, Appl
37	33.2	5.5	7001 1	US-08-258-261B-1 Sequence 1, Appl
38	33.2	5.5	7001 1	US-08-456-837-1 Sequence 1, Appl
39	33.2	5.5	7001 1	US-08-457-342-1 Sequence 1, Appl
40	33.2	5.5	7001 1	US-08-457-646A-1 Sequence 1, Appl
41	33.2	5.5	7001 1	US-08-458-076A-1 Sequence 1, Appl
42	33.2	5.5	7001 1	US-08-457-335A-1 Sequence 1, Appl
43	33.2	5.5	7001 1	US-08-761-258-6 Sequence 6, Appl
44	33.2	5.5	7001 1	US-08-729-214-1 Sequence 1, Appl
45	33.2	5.5	7001 2	US-08-977-306-6 Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 3670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232.463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935.313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZgpc-F15
US-08-232-463-14
Query Match 8.5%; Score 51.6; DB 1; Length 7218;

Best Local Similarity 3.8%; Pred. No. 8.5e-06;
Matches 15; Conservative 221; Mismatches 160; Indels 0; Gaps 0;

```
Oy 140 GCTCTGAGTCTTCCATTTTCACTTATTTCACTACATACCATCTTCCAGACGATATTC 199
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1058 GCTTCCGATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1117
Oy 200 AAGCTCTACATCCCAACCAATCCCATTTCAATCAACCACTTTAAAGCCTTGC 259
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1118 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1177
Oy 260 AAGGTACGTCTTCTTCTCAACGCGCCCGGAAATGACCAATCTCTCTCTCTTC 319
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1178 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1237
Oy 320 GTCAACCAATGTCAGTGGGTGAGATGATGATTAAGAAAGAAATTCATCAACCA 379
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1238 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1297
Oy 380 GTTCAACATGTTTCCCACTAAGGCTCCATTTCTCAAGCTGAATGACAGACCC 439
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1298 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1357
Oy 440 GATTCCTGAGCAACCAATTCAGCTCTCCAGACTAAGAACCATCTTCATCTTGTTC 499
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1358 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1417
Oy 500 GAGCTTCGAAATAGCGTTCGAAAGTCCGCTTACT 535
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1418 YYYYYYYYYYYYYYGGACCAATCTCTTATCT 1453
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 2

US-08-466-603-1/c

Sequence 1, Application US/08466603

Patent No. 5726018

GENERAL INFORMATION:

APPLICANT: Pasternack, Gary R.

APPLICANT: Kuhnajda, Francis P.

TITLE OF INVENTION: No. 5726018e1 Mammalian Protein Associated With

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS: Uncontrolled Cell Division

ADDRESS: Banner, Birch, McKie & Beckett

STREET: 1001 G Street, N.W.

CITY: Washington, D.C.

STATE: District of Columbia

COUNTRY: U.S.A.

ZIP: 20001

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/466,603

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/314,503

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Posorske Esq., Laurence H.

REGISTRATION NUMBER: 34,698

REFERENCE/DOCKET NUMBER: 1107,47218

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202 508-9153

TELEFAX: 202 508-9299

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1052 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 97..843
US-08-466-603-1

Query Match 6.6%; Score 40; DB 1; Length 1052;
Best Local Similarity 46.7%; Pred. No. 0.012;
Matches 127; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

```
Oy 59 AGAAGCGGAATGTAATCATCATGTTTGAATCTTGAATCAACCATCCGCTGAAT 118
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 879 ACAATAGGAATTTTCAAAAATAGTTATTCAGCTTAGTCATCATCTCTCCCTCATCTTC 820
Oy 119 GGTGTGTAATACCAATCATGCTCTGCAAGTCTTTTGCATATTTTCACTACATAC 178
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 819 AGTTCTGTTTGGCTTCTGACCCCTTCTTTCACCAAGCTTCTTTCATCTTCTC 760
Oy 179 CAGTCTCCAGCAAGTATTCAGCTCTCAATCCCAACCATCCCATTCAGTC 238
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 759 GTCACTACCTCTCCATGTTATACCTTTCATCTCTCTCTCTCTCACTACGTC 700
Oy 239 AAACATTAAAGCCTTCGCAAGTCACTGCTTCTTTCACCGCCGCGTGAATGAC 298
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 699 CTCCTTACACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 640
Oy 299 CAATTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 330
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 639 TTCAATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 608
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 3

US-08-314-503A-1/c

Sequence 1, Application US/08314503A

Patent No. 5734022

GENERAL INFORMATION:

APPLICANT: Pasternack, Gary R.

APPLICANT: Kuhnajda, Francis P.

TITLE OF INVENTION: No. 5734022e1 Mammalian Protein Associated With

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS: Uncontrolled Cell Division

ADDRESS: Banner, Birch, McKie & Beckett

STREET: 1001 G Street, N.W.

CITY: Washington, D.C.

STATE: District of Columbia

COUNTRY: U.S.A.

ZIP: 20001

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/314,503A

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Posorske Esq., Laurence H.

REGISTRATION NUMBER: 34,698

REFERENCE/DOCKET NUMBER: 1107,47218

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202 508-9153

TELEFAX: 202 508-9299

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1052 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA
ORIGINAL SOURCE: Homo sapiens
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 97..843
US-08-314-503A-1

Query Match 6.6%; Score 40; DB 1; Length 1052;
Best Local Similarity 46.7%; Pred. No. 0.012;
Matches 127; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 59 AGAAGCGGAATGTAATCATATGTTTGAATTTGAAATGCAACATGCCGCTGAATT 118
DB 879 ACAAATAGGAATTTTCAAAATAGTTATTCACATTAATCATATCTTCTCCATCTTC 820
QY 119 GGTGTCGTAATACCAATCATGCTGAGTCTTTTCCACTATTTCACCTCAATACC 178
DB 819 AGGTTCTGTTTTCGTTTGTGACCCCTTTCTTTCACCAAGCTCTTTCATCTTCCTC 760
QY 179 CAGTCTCCAGCAGCATTAATCAAGCTCTTCACATCCCAACCCATCCCATTCAGATC 238
DB 759 GTCATCTACTCTCCATCTGTAATTAACCTTTCATCTCTCTCTCTCACTCAAGTC 700
QY 239 AAACACTTTAAAGCTTTCGCAAGTCACTGTTCTTTTACCGCCGCGGGAATGAC 298
DB 699 CTCCTCTTACCT 640
QY 299 CAATTCCT 330
DB 639 TTCATCATCT 608

RESULT 4

US-08-468-066-1/c
Sequence 1, Application US/08468066
Patent No. 5756676

GENERAL INFORMATION:

APPLICANT: Pasternack, Gary R.
APPLICANT: Kunaide, Francis P.
TITLE OF INVENTION: No. 5756676el Mammalian Protein Associated With
TITLE OF INVENTION: Uncontrolled Cell Division
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie & Beckett
STREET: 1001 G Street, N.W.
CITY: Washington, D.C.
STATE: District of Columbia
COUNTRY: U.S.A.
ZIP: 20001

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,066
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/314,503
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Posorske Esq., Laurence H.
REGISTRATION NUMBER: 34,698
REFERENCE/DOCKET NUMBER: 1107.47218
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 508-9153
TELEFAX: 202 508-9299
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1052 base pairs
TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: 97..843

US-08-468-066-1

Query Match 6.6%; Score 40; DB 1; Length 1052;
Best Local Similarity 46.7%; Pred. No. 0.012;
Matches 127; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 59 AGAAGCGGAATGTAATCATATGTTTGAATTTGAAATGCAACATGCCGCTGAATT 118
DB 879 ACAAATAGGAATTTTCAAAATAGTTATTCACATTAATCATATCTTCTCCATCTTC 820
QY 119 GGTGTCGTAATACCAATCATGCTGAGTCTTTTCCACTATTTCACCTCAATACC 178
DB 819 AGGTTCTGTTTTCGTTTGTGACCCCTTTCTTTCACCAAGCTCTTTCATCTTCCTC 760
QY 179 CAGTCTCCAGCAGCATTAATCAAGCTCTTCACATCCCAACCCATCCCATTCAGATC 238
DB 759 GTCATCTACTCTCCATCTGTAATTAACCTTTCATCTCTCTCTCTCACTCAAGTC 700
QY 239 AAACACTTTAAAGCTTTCGCAAGTCACTGTTCTTTTACCGCCGCGGGAATGAC 298
DB 699 CTCCTCTTACCT 640
QY 299 CAATTCCT 330
DB 639 TTCATCATCT 608

RESULT 5

US-08-466-717-1/c
Sequence 1, Application US/08466717
Patent No. 5874234

GENERAL INFORMATION:

APPLICANT: Pasternack, Gary R.
APPLICANT: Kunaide, Francis P.
TITLE OF INVENTION: No. 5874234el Mammalian Protein Associated With
TITLE OF INVENTION: Uncontrolled Cell Division
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie & Beckett
STREET: 1001 G Street, N.W.
CITY: Washington, D.C.
STATE: District of Columbia
COUNTRY: U.S.A.
ZIP: 20001

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,717
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/314,503
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Posorske Esq., Laurence H.
REGISTRATION NUMBER: 34,698
REFERENCE/DOCKET NUMBER: 1107.47218
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 508-9153
TELEFAX: 202 508-9299
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

LENGTH: 1052 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 97..843
US-08-466-717-1

Query Match 6.6%; Score 40; DB 2; Length 1052;
Best Local Similarity 46.7%; Pred. No. 0.012;
Matches 127; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 59 AGAAGGGAATGTAATCATGTTTGAATTTGAAATCAACATGCGCTGAATT 118
DB 879 ACAATAGAAATTTTCAAAATAGTTATTCACATGTCATCTTCCCTCATCTTC 820
119 GGTGTGTAATACCAATCATGCTCTGACGCTTTTCCACTATTTTCACTCAATACC 178
DB 819 AGTTTCTCGTTTGGCTTCGACCCCTTTCTTTTCAACCAAGCTTTTCAATCTTC 760
QY 179 CAGTCTCCACAGCATGTAATCAAGCTTCACATCCGCCAACCATCCCATTAAGTC 238
DB 759 GTCATCTACCTCTCCATGTTATTAACCTTCTTCATCTCTCTCTCTCACTCAAGTC 700
QY 239 AAACATTTAAAGCTTTGCAAGGTCACTGTTCTTTTCAACGCGCGGTGAATGAC 298
DB 699 CTCCTCTTCACT 640
QY 299 CAATTCCT 330
DB 639 TTCATCATCT 608

RESULT 6

US-08-466-743-1/c
Sequence 1, Application US/08466743

GENERAL INFORMATION:
APPLICANT: Pasternack, Gary R.
APPLICANT: Khatada, Francis P.
TITLE OF INVENTION: No. 6040173el Mammalian Protein Associated With
TITLE OF INVENTION: Uncontrolled Cell Division
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie & Beckett
STREET: 1001 G Street, N.W.
CITY: Washington, D.C.
STATE: District of Columbia
COUNTRY: U.S.A.
ZIP: 20001

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,743
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/314,503
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Posorske Esq., Laurence H.
REGISTRATION NUMBER: 34,698
REFERENCE/DOCKET NUMBER: 1107.47218
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 508-9153
TELEFAX: 202 508-9299

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1052 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 97..843
US-08-466-743-1

Query Match 6.6%; Score 40; DB 3; Length 1052;
Best Local Similarity 46.7%; Pred. No. 0.012;
Matches 127; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 59 AGAAGGGAATGTAATCATGTTTGAATTTGAAATCAACATGCGCTGAATT 118
DB 879 ACAATAGAAATTTTCAAAATAGTTATTCACATGTCATCTTCCCTCATCTTC 820
QY 119 GGTGTGTAATACCAATCATGCTCTGACGCTTTTCCACTATTTTCACTCAATACC 178
DB 819 AGTTTCTCGTTTGGCTTCGACCCCTTTCTTTTCAACCAAGCTTTTCAATCTTC 760
QY 179 CAGTCTCCACAGCATGTAATCAAGCTTCACATCCGCCAACCATCCCATTAAGTC 238
DB 759 GTCATCTACCTCTCCATGTTATTAACCTTCTTCATCTCTCTCTCTCACTCAAGTC 700
QY 239 AAACATTTAAAGCTTTGCAAGGTCACTGTTCTTTTCAACGCGCGGTGAATGAC 298
DB 699 CTCCTCTTCACT 640
QY 299 CAATTCCT 330
DB 639 TTCATCATCT 608

RESULT 7

PCT-US95-12414-1/c
Sequence 1, Application PC/TU959512414

GENERAL INFORMATION:
APPLICANT: Pasternack, Gary R.
APPLICANT: Khatada, Francis P.
TITLE OF INVENTION: Novel Mammalian Protein Associated With
TITLE OF INVENTION: Uncontrolled Cell Division
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allegretti, Ltd.
STREET: 1001 G Street, N.W.
CITY: Washington, D.C.
STATE: District of Columbia
COUNTRY: U.S.A.
ZIP: 20001

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/12414
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/314,503
FILING DATE: 22-SEP-1994

ATTORNEY/AGENT INFORMATION:
NAME: Hoscheit Esq., Dale H.
REGISTRATION NUMBER: 19,090
REFERENCE/DOCKET NUMBER: 1107.51507
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 508-9153

337 GTGGCTCGAGATGGATTCTATAAAGACAGAATTCATCCAAAGTTCAACATGTTTTTC 396

Db 65 YNVSNNVCATYTYGTYAAYTYTGTAAAYTAAYTAAYTAAYTYTGYTC 6
QY 397 CCACT 401
Db 5 YCYCY 1

RESULT 10
US-08-766-738-2/c
Sequence 2, Application US/08766738
Patent No. 5916749

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: NOVEL HUMAN PHOSPHATASE INHIBITOR PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/766,738
FILING DATE: Herewith
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0177 US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166

TELEX:
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 966 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear

IMMEDIATE SOURCE:
LIBRARY: Consensus

CLONE: 1813361
US-08-766-738-2

Query Match 6.2%; Score 37.8; DB 2; Length 966;
Best Local Similarity 50.8%; Pred. No. 0.054;
Matches 90; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 150 TCTTTTCCATATTTTCACTTCAACATACCAAGTCTTCCAGACGATTTCAAGCTCTCA 209
Db 836 TCTCTTCAATCAATCTGTTTCTCTCTCTTTTCAACCTTTCCACCTTCTCTCTCT 777
QY 210 CATCCCCCAACCAATCCCATTCATCAAGTCAAAACATTTAAAGCTTCCGCAAGTCACTG 269
Db 776 TCATCTCATCTCATCT 717
QY 270 TCTTTTCTTTCACCGCGCGCGGGAATGACCAATTCCTCTCTCTCTCTCTCTCTCTCA 326
Db 716 ACTTCATCTGTCCTCTCAATCCCTTCTCAATCTTCAATCTTCAATCTTCTCTCA 660

RESULT 11

US-09-262-610-2/c
Sequence 2, Application US/09262610
Patent No. 6428949

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: NOVEL HUMAN PHOSPHATASE INHIBITOR PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/262,610
FILING DATE:
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/766,738
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0177 US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166

TELEX:
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 966 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear

IMMEDIATE SOURCE:
LIBRARY: Consensus

CLONE: 1813361
US-09-262-610-2

Query Match 6.2%; Score 37.8; DB 4; Length 966;
Best Local Similarity 50.8%; Pred. No. 0.054;
Matches 90; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 150 TCTTTTCCATATTTTCACTTCAACATACCAAGTCTTCCAGACGATTTCAAGCTCTCA 209
Db 836 TCTCTTCAATCAATCTGTTTCTCTCTCTCTTTTCAACCTTTCCACCTTCTCTCTCT 777
QY 210 CATCCCCCAACCAATCCCATTCATCAAGTCAAAACATTTAAAGCTTCCGCAAGTCACTG 269
Db 776 TCATCTCATCTCATCT 717
QY 270 TCTTTTCTTTCACCGCGCGCGGGAATGACCAATTCCTCTCTCTCTCTCTCTCTCTCA 326
Db 716 ACTTCATCTGTCCTCTCAATCCCTTCTCAATCTTCAATCTTCAATCTTCTCTCA 660

RESULT 12

US-09-178-252-26
Sequence 26, Application US/09178252
Patent No. 6218188

GENERAL INFORMATION:
APPLICANT: Cardneau, Guy A.
APPLICANT: Stelman, Steven J.
APPLICANT: Narva, Kenneth E.
TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins
FILE REFERENCE: WA-714XC2

CURRENT APPLICATION NUMBER: US/09/178,252
CURRENT FILING DATE: 1998-10-23
EARLIER APPLICATION NUMBER: 60/065,215
EARLIER FILING DATE: 1997-11-12
EARLIER APPLICATION NUMBER: 60/076,445
EARLIER FILING DATE: 1998-03-02
NUMBER OF SEQ ID NOS: 27
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO: 26
LENGTH: 1965
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic B.t. toxin gene
US-09-178-252-26

Query Match
Best Local Similarity 52.3%; Score 36.2; DB 4; Length 1965;
Pred. No. 0.23; Mismatches 0; Gaps 0;
Matches 80; Conservative 0; Indels 0; Gaps 0;

176 ACCGAGCTCTCCGAGCAGATTTCAAGCTCTCCATCCCGCCAGCCATCCCATTTCAA 235
1336 ACCAAGCGCGCAACACCTGCGCCGAGACCTCACCACGAGAGAGTCCGCTGGAG 1395
QY 236 GTCAACACTTTAAAGCTTCCGAAAGTCACTGTCTTCTTCCACCGCCGCTGAAT 295
DB 1396 AACAAACACTTCAACCTCTCTCCGACGTGACCTTCTCGCTTCAACACCCAGGGC 1455
QY 296 GACCAATCT 328
DB 1456 GGCCCACTCGCAGCGCTGGGGTTCGTCCGACC 1488

RESULT 13

US-08-574-959A-8/c
Sequence 8, Application US/08574959A
Patent No. 5962224
GENERAL INFORMATION:
APPLICANT: Jaekyoon Shin, Inseil Joung, Ratna K. Vadlamudi
APPLICANT: and Jack L. Strominger
TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
TITLE OF INVENTION: AND USES THEREFOR
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/574,959A
FILING DATE: 19-DEC-95
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: DFN-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 3211 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:

NAME/KEY: CDS
LOCATION: 439..3157
US-08-574-959A-8

Query Match
Best Local Similarity 49.5%; Score 35.6; DB 2; Length 3211;
Pred. No. 0.44; Mismatches 94; Indels 0; Gaps 0;
Matches 92; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 150 TCTTTTCACATTTTACCTCAATCCCACTCTTCCGACGATTTCAAGCTCTCA 209
-DB 2609 TCT 2550
QY 210 CATCCCCCAACCAATCCCATTTCAAGTCAAACTTTAAAGCTTGGCAAGTCACTG 269
DB 2549 TCTTCT 2490
QY 270 TCTTCT 329
DB 2489 TCTTCT 2430
QY 330 TGTGCC 335
DB 2429 TCTTCC 2424

RESULT 14

US-09-357-014-8/c
Sequence 8, Application US/09357014
Patent No. 6291645
GENERAL INFORMATION:
APPLICANT: Jaekyoon Shin, Inseil Joung, Ratna K. Vadlamudi
APPLICANT: and Jack L. Strominger
TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
TITLE OF INVENTION: AND USES THEREFOR
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/357,014
FILING DATE: 19-Jul-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/574,959
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: DFN-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 3211 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 439..3157
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-357-014-8

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2003, 06:37:48 ; Search time 109 Seconds
(without alignments)
8293.844 Million cell updates/sec

Title: US-10-021-323-13
Perfect score: 609
Sequence: 1 ggaatgaatcaactttt.....tgaagcttacaataaag 609

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1055720 seqs, 742p2436 residues
al number of hits satisfying chosen parameters: 211440

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications NA.*
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq.*
11: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq.*
13: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq.*
14: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	82.4	13.5	279	10	US-09-878-574-5466 Sequence 5466, App
C 2	81.2	13.3	381	10	US-09-770-791-166 Sequence 166, App
C 3	78.6	12.9	244	10	US-09-878-574-10284 Sequence 10284, A
C 4	55.6	9.1	576	9	US-09-938-842A-2017 Sequence 2017, App
C 5	53.6	8.8	510	9	US-09-938-842A-1597 Sequence 1597, App
C 6	40.8	6.7	671	9	US-10-184-644-346 Sequence 346, App
C 7	40.8	6.7	671	9	US-10-184-644-346 Sequence 346, App
C 8	40.4	6.6	520	9	US-10-184-644-332 Sequence 332, App
C 9	40.4	6.6	520	9	US-10-184-644-332 Sequence 332, App
C 10	39.6	6.5	423	9	US-09-938-842A-2157 Sequence 2157, App
C 11	38.8	6.4	7183	9	US-10-198-846-11576 Sequence 11576, A
C 12	38.4	6.3	369	10	US-09-770-791-382 Sequence 382, App
C 13	37.8	6.2	966	9	US-10-213-700-2 Sequence 2, App1
C 14	37.6	6.2	966	9	US-10-213-700-2 Sequence 2, App1
C 15	36.2	5.9	1965	10	US-09-826-660-26- Sequence 26, App1
C 16	35.8	5.9	486	9	US-09-938-842A-2157 Sequence 2157, App
C 17	35.8	5.9	487	10	US-09-864-761-10682 Sequence 10682, A
C 18	35.8	5.9	515	10	US-09-864-761-30248 Sequence 30248, A
C 19	35.8	5.9	174566	9	US-10-020-141-1 Sequence 1, App1

C 20	35.6	5.8	368	10	US-09-983-965-73 Sequence 73, App1
C 21	35.2	5.8	1434	9	US-09-738-626-283 Sequence 283, App
C 22	35.2	5.8	3309400	9	US-09-738-626-1 Sequence 1, App1
C 23	34.4	5.6	856	10	US-09-967-347-15 Sequence 15, App1
C 24	34.4	5.6	3397	9	US-10-198-846-11555 Sequence 11555, A
C 25	34.4	5.6	6688	9	US-09-813-453A-72 Sequence 72, App1
C 26	34.4	5.6	12241	12	US-10-033-190-5 Sequence 5, App1
C 27	34.4	5.6	13737	9	US-10-074-279-10 Sequence 10, App1
C 28	34.4	5.6	14446	9	US-09-810-861B-4 Sequence 4, App1
C 29	34.2	5.6	748	10	US-09-770-149-17 Sequence 17, App1
C 30	34.2	5.6	1425	9	US-09-738-626-959 Sequence 959, App
C 31	34.2	5.6	1842	9	US-09-938-842A-1028 Sequence 1028, App
C 32	34	5.6	470	9	US-09-918-995-3480 Sequence 3480, A
C 33	33.8	5.6	391	9	US-09-918-995-5449 Sequence 5449, App
C 34	33.8	5.6	449	9	US-09-918-995-11791 Sequence 11791, A
C 35	33.8	5.6	455	9	US-09-918-995-15644 Sequence 15644, A
C 36	33.8	5.6	459	9	US-09-918-995-15644 Sequence 15644, A
C 37	33.8	5.6	465	9	US-09-918-995-3284 Sequence 3284, App
C 38	33.8	5.6	478	9	US-09-918-995-14016 Sequence 14016, A
C 39	33.8	5.6	486	9	US-09-918-995-9799 Sequence 9799, App
C 40	33.8	5.6	1120	10	US-09-969-708-320 Sequence 320, App
C 41	33.8	5.6	1261	9	US-10-119-926-59 Sequence 59, App1
C 42	33.6	5.5	350	10	US-09-783-590-7153 Sequence 7153, App
C 43	33.6	5.5	766	10	US-09-864-761-19608 Sequence 19608, A
C 44	33.6	5.5	1944	10	US-09-864-761-2825 Sequence 2825, App
C 45	33.6	5.5	2082	10	US-09-846-996A-2 Sequence 2, App1

ALIGNMENTS

RESULT 1	US-09-878-574-5466/c	Application US/09878574
Sequence 5466, App	Patent No. US2002010548A1	
GENERAL INFORMATION:		
APPLICANT:	Byrum, Joseph R.	
APPLICANT:	La Rosa, Thomas J.	
APPLICANT:	Thompson, Michael D.	
TITLE OF INVENTION:	Nucleic Acid Molecules and Other Molecules Associated with	
TITLE OF INVENTION:	Plants	
FILE REFERENCE:	38-21(15401)B	
CURRENT APPLICATION NUMBER:	US/09/878, 574	
CURRENT FILING DATE:	2001-12-21	
PRIOR APPLICATION NUMBER:	09/333, 535	
PRIOR FILING DATE:	1999-06-14	
NUMBER OF SEQ ID NOS:	15775	
SEQ ID NO 5466		
LENGTH:	279	
TYPE:	DNA	
ORGANISM:	Glycine max	
OTHER INFORMATION:	Clone ID: 701097008H1	
US-09-878-574-5466		
Query Match	13.5%	Score 82.4; DB 10; Length 279;
Best Local Similarity	62.5%	Pred. No. 4.8e-16;
Matches 145; Conservative	0; Mismatches 86;	Indels 1; Gaps 1;
Cy 345	GAGATGATTCATTAAGAAACAAGATTCATTCAGGTCAGGATGTTCCCACTAAG	404
Db 279	GATATGATGTTGTAAGAAACAAGATTCATTCAGGTCAGGATGTTCCCACTAAG	221
Cy 405	GAGTCCATTCATTAAGGCTGAATGACAGCCGATTCCTGAGCAACAATTCAGC	464
Db 220	GATTCACATTCCTATGCTATGATGGAATTCCTGATTCGAGCAACAATTCAGT	161
Cy 465	TCCATCAGATTAAGCAAGCATTCATTCCTGAGCTTCGCAATTCAGCTTCAG	524
Db 160	TCCATCAGGCTCAGCATTCGCTGATTCAGCTTCGCAATTCAGCTTCAG	101
Cy 525	TCCGCTTACTAAGGGGACATTTTACCAACAAGATGATGATCACCAT	576
Db 100	TCAATGCTGTTAGGAGACACATATTCGCTATTAACATGATTTAAAGCT	49

RESULT 2

US-09-770-791-166
Sequence 166, Application US/09770791
Patent No. US20020062014A1
GENERAL INFORMATION:
APPLICANT: Goriach, Jörn
APPLICANT: An, Yong-Qiang
APPLICANT: Hamilton, Carol M.
APPLICANT: Price, Jennifer L.
APPLICANT: Raines, Tracy M.
APPLICANT: Yu, Yang
APPLICANT: Rameaka, Joshua G.
APPLICANT: Page, Amy
APPLICANT: Mathew, Abraham V.
APPLICANT: Ledford, Brooke L.
APPLICANT: Moessner, Jeffrey P.
APPLICANT: Haas, William David
APPLICANT: Garcia, Carlos A.
APPLICANT: Kricker, Maja
APPLICANT: Slader, Ted
APPLICANT: Davis, Keith R.
APPLICANT: Allen, Keith
APPLICANT: Hoffman, Neil
APPLICANT: Hurlan, Patrick
TITLE OF INVENTION: Expressed Sequences of Arabidopsis
FILE REFERENCE: 2029 (PARA-018PRV)
CURRENT APPLICATION NUMBER: US/09/770,791
CURRENT FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/178,480
PRIOR FILING DATE: 2000-01-27
NUMBER OF SEQ ID NOS: 999
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 166
LENGTH: 381
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-770-791-166

Query Match 13.3%; Score 81.2; DB 10; Length 381;

Best Local Similarity 61.2%; Pred. No. 1,4e-15;
Matches 131; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 68 ATGTAAATCATGTTTGAATCTTGAATAACACATGCGCTGAATTGGTGTGTA 127
167 ATTTAATCATGTTTCTTGAATCTTGAATCTTGAATCAAAACATCAAGATTCTGTATG 226
WY 128 ATACCAATCATGCTCTGCGAGTCTTTTCCACTATTTCACCTCACATACCCAGTCTCC 187
227 AACTCGATCATCTCCCAACATCCCAAGCTTTCCTCTCCCAAAACCTTACGCTTC 286
QY 188 CAGCAGTATTCAGAGCTCTTCACATCCCAACCCATCCCAATTCAGTCAAACTTT 247
287 CACACATCAACGAGCTCTCCCTCGTGAATATTAACCATTCCTTCATCATGAACATTT 346
DB 248 AAAAGCCTTCGACAGGTCATGCTCTTCTTCTCA 281
347 AAACGCTTCGCAATGCTTCATCATATATACGA 380

RESULT 3

US-09-878-574-10284/c
Sequence 10284, Application US/09878574
Patent No. US2002010548A1
GENERAL INFORMATION:
APPLICANT: Byrum, Joseph R.
APPLICANT: La Rosa, Thomas J.
APPLICANT: Thompson, Michael D.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21 (15401) B

CURRENT APPLICATION NUMBER: US/09/878,574

CURRENT FILING DATE: 2001-12-21

PRIOR APPLICATION NUMBER: 09/333,535

PRIOR FILING DATE: 1999-06-14

NUMBER OF SEQ ID NOS: 15775

SEQ ID NO 10284

LENGTH: 244

TYPE: DNA

ORGANISM: Glycine max

OTHER INFORMATION: Clone ID: 701103505H1
US-09-878-574-10284

Query Match 12.9%; Score 78.6; DB 10; Length 244;

Best Local Similarity 60.6%; Pred. No. 7.2e-15;
Matches 129; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 364 ACAAGATTGATCAAGTTCAAAATGATGTTTCCCACTAAGGCTCAATTCTTCAAGC 423
244 ACAAGAAATCACTGAACCAAGACTTTTCTTCCCAAGTGATTCCAACTCTTCTATGC 185
QY 424 TGAATTGACAGACCCGATTCCTGGAGCAACCAATTCAGCTCCCTCAGACTAAGAGC 483
184 TATATTGGGAATTCGCCGTCATGTGAACATGATTCATGTTCTCAAGGCTCAACAT 125
QY 484 CATCTCCATTTCTGTGAGCTTCTCGAATACGGTTGCAAGTCGTCTTAAGAGGAGG 543
124 CCGCGCATTCACGTCTCACTTCTCAAAAATACGCTTCAAGCAGTTGCTGTAGAGAC 65
QY 544 ACATTTTACCACCAAGAGATATGATCAACAT 576
DB 64 AGTTATTGCAATAACGCAATTTAAAGCCT 32

RESULT 4

US-09-938-842A-2017/c
Sequence 2017, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: S01P1300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 2017
LENGTH: 576
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-2017

Query Match 9.1%; Score 55.6; DB 9; Length 576;

Best Local Similarity 49.7%; Pred. No. 2.8e-07;
Matches 170; Conservative 0; Mismatches 169; Indels 3; Gaps 1;

QY 74 CATCATGTTTGAATTTCTGAAAATCAACATGCGGTGAATTGGTGTGAATACCA 133
549 CATCATGTTTGAATTTCAAAAAGTCAACGACCATTTATTCGGTCAACGGAAC 490
QY 134 AATCATGCTCCGAGAGTCTTTCACACTATTTTCACCTACATACCAAGGCTCCAGAC 193
489 GATCATCTTCTCCACTCTTTCATTTCTCCACTTCAGTAAACAAAGTTTTCGAAC 430
DB 194 GTATTCAGCTCCTCATCATCCCAACCATCCCATTCATCAAGTCAAAACATTAAAGC 253

GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Krieps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
TITLE OF INVENTION: SAME, AND METHODS OF USE
FILE REFERENCE: SRIPI300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 592
LENGTH: 423
TYPE: DNA
ORGANISM: Arabidopsis thaliana
-09-938-842A-592

Query Match 6.5%; Score 39.6; DB 9; Length 423;
Best Local Similarity 69.2%; Pred. No. 0.031;
Matches 54; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 253 CCTCGCAGAGTCACTGTCTTCTTCAACCGCGCGGTGAATGACCAATTCCTCCT 312
DB 349 CCTCGCAGAGTCACTGTCTTCTTCAACCGCGCGGTGAATGACCAATTCCTCCT 290

QY 313 CCTCTTGTACACCAT 330
DB 289 CCTCATCTCATCATCGT 272

RESULT 11

US-10-198-846-11576/c
Sequence 11576, Application US/10198846
Publication No. US2003009974A1
GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steilmann, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
TITLE OF INVENTION: THERAPY OF BREAST CANCER
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198,846
CURRENT FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/306,220
PRIOR FILING DATE: 2001-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11576
LENGTH: 7183
TYPE: DNA
ORGANISM: Homo sapiens
US-10-198-846-11576

Query Match 6.4%; Score 38.8; DB 9; Length 7183;
Best Local Similarity 48.2%; Pred. No. 0.34;
Matches 109; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 244 CTTTAAAGCCTTGGAGGTCAGTCTTCTTCAACCGCGCGGTGAATGACCAATT 303
DB 4443 CTTAATCATGCGGTGATCATCATCTTCAATCATCTTCTTCTTCTTCTTCTT 4384

QY 304 CCTCCTCTCTCTTCTGTCACCAATGTCGCGGTGTCGAGATGATTCATAAAGA 363
DB 4383 CTTGCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 4324

QY 364 ACAAGATTCATCCAGATTCGAACATGTTTTCCACATGAGGCTCCATTTCTTCAAGC 423
DB 4323 TGCATACCTGTGTCATGTCAGGCAATCCATGCTGCTGATTTCTTCAATGACGAGAT 4264
QY 424 TGAATGACAGACCGGATTTCTGAGCAACCAATTCAGTCTC 469
DB 4263 AGCATGCTTGTGTGTCGTACAGGTCAAGATGTTCCGTTTCAAC 4218

RESULT 12

US-09-770-791-382/c
Sequence 382, Application US/09770791
Patent No. US2003006201A1
GENERAL INFORMATION:
APPLICANT: Goriach, Jörn
APPLICANT: An, Yong-Qiang
APPLICANT: Hamilton, Carol M.
APPLICANT: Price, Jennifer L.
APPLICANT: Raines, Tracy M.
APPLICANT: Yu, Yang
APPLICANT: Rameaka, Joshua G.
APPLICANT: Page, Amy
APPLICANT: Mathew, Abraham V.
APPLICANT: Ledford, Brooke L.
APPLICANT: Woessner, Jeffrey P.
APPLICANT: Haas, William David
APPLICANT: Garcia, Carlos A.
APPLICANT: Krieger, Maja
APPLICANT: Slader, Ted
APPLICANT: Davis, Keith R.
APPLICANT: Allen, Keith
APPLICANT: Hoffman, Neil
APPLICANT: Hurdan, Patrick
TITLE OF INVENTION: Expressed Sequences of Arabidopsis
TITLE OF INVENTION: thaliana
FILE REFERENCE: 2029 (PAPA-018PRV)
CURRENT APPLICATION NUMBER: US/09/770,791
CURRENT FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/178,480
PRIOR FILING DATE: 2000-01-27
NUMBER OF SEQ ID NOS: 999
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 382
LENGTH: 369
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-770-791-382

Query Match 6.3%; Score 38.4; DB 10; Length 369;
Best Local Similarity 49.5%; Pred. No. 0.068;
Matches 99; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 56 TCAAGACCGGAATGATCATGTTTGAATCTTGAATAATCAACATGCCGCTGA 115
DB 299 TCAATAAAGAGTGTCTCATGATTAACAACTGTTAAATGATTTTCAATCTT 240

QY 116 ATTGCTGTGATTAACCAATCATGCTCTGAGTCTTTTCAATTTTCACTACAT 175
DB 239 GTTCCATTAATGACTTAATCATTAACCAATTTGTTAGTAGATCTTGTATAA 180

QY 176 ACCAGTCTTCCAGACGATTAATCAAGCTCTTCAATCCCAACCAATCCCATYCA 235
DB 179 CCTAAGATTTGCAAAAGCTCTGTAACTCTAATGATCAATGAACATCTTGTGTTTC 120

QY 236 GTCAAACTTTAAAGCTT 255
DB 119 GTCAAACTATCAAAAGCTT 100

RESULT 13
US-10-213-700-2/c
Sequence 2, Application US/10213700
Publication No. US20030022332A1

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
Gail, Surya K.
TITLE OF INVENTION: NOVEL HUMAN PHOSPHATASE INHIBITOR PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/213,700
FILING DATE: 06-Aug-2002
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/766,738
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0177 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 966 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Consensus
CLONE: 1813161
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-213-700-2
Query Match 6.2%; Score 37.8; DB 9; Length 966;
Best Local Similarity 50.8%; Pred. No. 0.2;
Matches 90; Conservative 0; Mismatches 87; Indels 0; Gaps 0;
Db 150 TCTTTTCACATATTTTCACCTCAATACCCAGTCTTCCAGACGATTCAGCTCTCA 209
836 TCTCCTTCATCATCTGTTCTCTCTCTTTCACCTTCCCACTTCTTCTCTCT 777
Qy 210 CATCCGCCCAACCCATCCCATTCAGTCAACACTTTAAAGCTTCGAGGTCTG 269
Db 776 TCATCTCATCTCATCTTCATCTTCTTCAAGTCCAAATCTTCTTCTCTCACTG 717
Qy 270 TCTTCTTCTTACCGCGCGCGGTGAATGACCAATTCCTCCCTCCCTTGGTACCA 326
Db 716 ACTTCATCTGTCGTCCTCATCTCCCTTCACTTTCATCTTCACTTCTTCACTCA 660
RESULT 14
US-10-123-155-10/c
Sequence 10, Application US/10123155
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Neutreen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Geo, Wei-Qiang
APPLICANT: Gerlitsen, Mary E.

APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gutney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C30
CURRENT APPLICATION NUMBER: US/10/123,155
CURRENT FILING DATE: 2002-04-15
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 10
LENGTH: 594
TYPE: PRT
ORGANISM: Homo Sapien
US-10-123-155-10
Query Match 6.2%; Score 37.6; DB 9; Length 594;
Best Local Similarity 10.4%; Pred. No. 0.17;
Matches 31; Conservative 116; Mismatches 150; Indels 0; Gaps 0;
Db 132 CAATCATCTCTCGAGCTTTTCCACATTTTTCACCATACCCAGTCTCCAGC 191
556 MM.M.CBT.STT.YA.M.YT.S.S.S.SYSYSYS.S.S.SDSYSYA.SYSYS.S.SMS 497
Qy 192 ACGTATTCAGCTCTCATCATCCCAACCCATCCCATTCAGTCAACACTTTAA 251
Db 496 SYSYSSDDY.CYCCYRHCSDSYSYVY.CRCGYR.SVSRDCHYSCCSDYCY 437
Qy 252 GCGTTCGACAGTCTCTTCTTCTTACCGCGCGCGGTGAATGACCAATCTCTCC 311
Db 436 YSYSRYSYSYSYSYSYTDYCSYRCCYYSYSYSSSSSSSSSSSSSSSS 377
Qy 312 TCTCTCTGTCACCAACATGTCGCGGTGATGATGATTCATTAAGAACAGAT 371
Db 376 YVTSNYC.T.CC...T.MCABGCTTTTTTTTTT.HSCC.SA.A.M.YC.A.S 317
Qy 372 TCATCCAGTCAACATGTTTCCCACTAAGGCTCCCATTTCTTCAAGCTGAT 428
Db 316 YSYSYS.SSS.S.SYMR.HRA.SHYTRS.S.MCY.YM.Y.YY.YSYSYCSRT 260
RESULT 15
US-09-826-660-26
Sequence 26, Application US/09826660
Patent No. US20010026540A1
GENERAL INFORMATION:
APPLICANT: Cardineau, Guy A.
APPLICANT: Steiman, Steven J.
APPLICANT: Narva, Kenneth E.
TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins
FILE REFERENCE: MA-714XC2D1
CURRENT APPLICATION NUMBER: US/09/826,660
CURRENT FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 09/178,252
PRIOR FILING DATE: 1998-10-23
PRIOR APPLICATION NUMBER: 60/065,215
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/076,445
PRIOR FILING DATE: 1998-03-02
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 26
LENGTH: 1965
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURES:

OTHER INFORMATION: Synthetic B.t. toxin gene
US-09-826-660-26

Query Match	5.9%	Score 36.2;	DB 10;	Length 1965;
Best Local Similarity	52.3%;	Pred. No. 1;		
Matches 80;	Conservative	0;	Mismatches 73;	Indels 0;
				Gaps 0;

QY 176 ACCGAGTTCGCCAGCAAGTAATCAAGTCCCAACATCCCCCAACCATCCCAATCAA 235
Db 1336 ACCAAGCGCGCCCAACCTGCGCCGCGAGGACCTTCACACACACCGAGAGACTCCCGCTGGAG 1395
QY 236 GTCAACACCTTTAAAGCCTTCGCAAGGTCACTGTCCTTCTTCAACCGCGCGCATGAT 295
Db 1336 AACACAACTTCACACTCCTCTCCCACTGACCTTCTCCGCTTCACACACCAACCAAGGC 1455
QY 296 GACCAATTCCTCTCCTCCTCTCTTCGTCAACAC 328
Db 1456 GCGCCACTCGCAGCGCTGGGGTGTCTCCGACGC 1488

```

rch completed: June 29, 2003, 07:24:55
time : 112 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2003, 06:37:18 ; Search time 1076 Seconds
(without alignments)
9166.405 Million cell updates/sec

Title: US-10-021-323-13

Perfect score: 609
Sequence: 1 ggtaactgaatacaacttctt.....tgaagctncaataaagg 609

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 16154066 seqs, 8b97743b76 residues

al number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estrom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_hiv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	200.8	33.0	718	13	BM436482 VVA004H01
C 2	177.4	29.1	572	10	BE240668 EST404717
C 3	146.6	24.1	535	13	BI426522 BSG04C05
C 4	145.6	23.9	496	10	AM156763 BSG04D10.Y
C 5	145.4	23.9	516	13	BI425377 BAF28612
C 6	142.6	23.4	567	10	AM171688 N100582e

7	140.4	23.1	667	10	BE588834
C 8	134.8	22.1	433	10	AM830375
C 9	111.6	18.3	245	10	BE240144
C 10	107.2	17.6	526	12	BF424862
C 11	101.2	16.6	590	10	AV827067
C 12	100.6	16.5	494	9	AT779002
C 13	94.8	15.6	382	12	BE999195
C 14	92	15.1	395	12	BE997593
C 15	88	14.4	499	10	AM266833
C 16	85.4	14.0	756	17	BE677796
C 17	84.8	13.9	734	14	BQ986853
C 18	79.6	13.1	455	14	T76748
C 19	78	12.8	441	10	AT797137
C 20	75.2	12.3	564	14	BQ987606
C 21	74.2	12.2	421	12	BE268779
C 22	73.4	12.1	655	14	BQ407687
C 23	73.4	12.1	658	14	BQ412536
C 24	67.4	11.1	447	9	AT993426
C 25	65.8	10.8	530	9	AJ319937
C 26	65.6	10.8	553	10	AM586577
C 27	65.6	10.8	555	10	BE124592
C 28	65.6	10.8	684	12	BE634668
C 29	62.6	10.3	537	9	AT774144
C 30	61.6	10.1	555	14	BQ079515
C 31	59.8	9.8	437	14	BM731250
C 32	59.8	9.8	520	12	BE047423
C 33	59.8	9.8	650	13	BI969420
C 34	59	9.7	309	14	N37193
C 35	58.6	9.6	471	10	BE555981
C 36	58.2	9.6	327	12	BE777638
C 37	58.2	9.6	480	12	BE673834
C 38	58.2	9.6	534	12	BM317786
C 39	58.2	9.6	573	13	BM371103
C 40	57.6	9.5	541	10	BE436891
C 41	57.6	9.5	548	10	AM030305
C 42	57.6	9.5	659	17	BE588128
C 43	57.6	9.5	825	12	BE126826
C 44	56.8	9.3	502	10	AM985455
C 45	56.6	9.3	912	12	BE342900

ALIGNMENTS

RESULT 1
LOCUS BM436482 718 bp mRNA linear EST 31-JAN-2002
DEFINITION VVA004H01.52547 An expressed sequence tag database for abiotic stressed leaves of Vitis vinifera var. Chardonnay Vitis vinifera CDNA clone VVA004H01.5, mRNA sequence.

ACCESSION BM436482
VERSION BM436482.1 GI:18458204

KEYWORDS
SOURCE Vitis vinifera.
ORGANISM Vitis vinifera.

REFERENCE 1 (bases 1 to 718)
AUTHORS Cramer, G.R. and Cushman, J.C.
TITLE An expressed sequence tag database for abiotic stressed leaves of Vitis vinifera var. Chardonnay
JOURNAL Unpublished (2002)
COMMENT Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
PCR Primers
FORWARD: T3 20mer
BACKWARD: T7 21mer

Plate: 004 row: H column: 01
 Seq primer: T3 20mer
 High quality sequence step: 718.
 Location/Qualifiers

FEATURES

source

1. 718
 /organism="Vitis vinifera"
 /db_xref="taxon:29760"
 /clone="VVA004H01"
 /clone_1ib="An expressed sequence tag database for abiotic
 stressed leaves of Vitis vinifera var. Chardonnay"
 /tissue_type="leaf"
 /dev_stage="juvenile and adult"
 /note="Vector: Lambda Uni-Zap XR. Bluescript SK-; Site 1:
 EcoRI; Site 2: XhoI; library construction was performed
 according to Stratagene's recommended protocol for the
 Lambda UniZapXR vector and cDNA synthesis kit."

BASE COUNT 185 a 146 c 189 g 198 t

ORIGIN

Query Match 33.0%; Score 200.8; DB 13; Length 718;
 Best Local Similarity 66.8%; Pred. No. 9.1e-50;
 Matches 318; Conservative 0; Mismatches 152; Indels 6; Gaps 2;

QY 70 GTAACATCATGTTTGAATTTCTGAAATTCACACCATGCGCTGAATTGCTGTAAT 129
 DB 529 GAGACATCATGTTTCTGAACTCTCAAAATCAAGAACCCATGAGTTGTTCAATCA 470
 QY 130 ACCAAATCATGCTCTGCAAGTCTTTTCCATATTTTCACTGACATACCAAGCTTTCCA 189
 DB 469 TCTGATCATGCTCTGCAATGACCTCCACCATTTCTCCCATTTCCAGTCTGAGCA 410
 QY 190 GCAAGATTCAGCTCTCTCATATCCCAACCCATTCCTCATTAAGTCAAAAGCTTTAA 249
 DB 409 GAGCGCTGCAAGCTCATACATGATGAAACCATCTCGTTCAAGTCAAAACCTGTA 350
 QY 250 AAGCCTTGCAGAGTCACTGCTCTTCTTCTTCAACGCGCGGGAATGACCAATTCCTCC 309
 DB 349 AAGCTTCTGAGTCAACCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 290
 QY 310 CCTCTCTTCTGTCACCAACATGTCAGAGGTTTCAGATGATTAAGAAAGCAAGA 369
 DB 289 CCTCGGCTCTGCT 233
 QY 370 ATTATCTCAAGTCAAAACATGTTTTCCTCACTAAGGCTTCAATTTCTTCAAGGCTAAT 429
 DB 232 ACTCATTAATAATGAGAGGTGTTTCCCAAGAGGCTCTCAAGCTATCCAGGCTTACT 173
 QY 430 GGACAGACCCGATTTCTGAGAGCAACAAATTCAGTCTCTCCAGACTAAGCAAGCCATTC 489
 DB 172 GCAACA--CGAACTCTCTCAAGAGCAAGTTAGCTTCCCAAGGCTCAACAGGCCATCAC 116
 QY 490 CATTTCTTGTGAGCTTCTGAAATACCGGTTGCAAGTCTGCTTACTTAAGGGGGAGC 545
 DB 115 CATTTGCTGCAAGCTTGTGAAAATCTCTGACAGGTCATTTGAGCTCAAGGAAGAC 60

RESULT 2

BE240668/c 572 bp mRNA linear EST 12-JUL-2000
 LOCUS BE240668/c MRP- Medicago truncatula cDNA clone pmMRP-47M18, mRNA
 DEFINITION sequence.

ACCESSION BE240668
 VERSION BE240668.1 GI:9056852
 KEYWORDS EST.
 SOURCE barrel medic.
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot;
 Rosidae; eurosoid I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
 Medicago.

REFERENCE 1 (bases 1 to 572)
 Harrison,M.J., Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S.,
 Holt,I.E., Cho,J. and Fraser,C.M.

TITLE ESTs from phosphate-starved roots of Medicago truncatula
 JOURNAL Unpublished (2000)
 COMMENT Contact: Maria J. Harrison
 Plant Biology Division
 The Samuel Roberts Noble Foundation
 2510 Sam Noble Parkway, Ardmore, OK 73401, USA
 Tel: 580-223-5810
 Fax: 580-221-7380
 Email: mjharrison@noble.org
 The Samuel Noble Roberts Foundation: N266535e
 TIGR sequence name: MTHAV81TX
 More information is available at
 http://chryslie.tamu.edu/medicago
 Seq primer: Skmod (CTA GAA CTA gta gat CC).

FEATURES

source

1. 572

/organism="Medicago truncatula"

/cultivar="A17"

/db_xref="taxon:3980"

/clone="pmMRP-47M18"

/clone_1ib="MRP"

/tissue_type="roots"

/dev_stage="phosphate-starved"

/lab_host="XhoI/R"

/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
 XhoI; At the trifoliolate stage, M. truncatula plants were
 transplanted to phosphate-free sand and grown for a
 further 30 days. During this period, they were fertilized
 twice weekly with 1/2 Hoaglands solutions containing 20mM
 potassium phosphate. cDNA was prepared from polyA+
 enriched RNA. The cDNA was directionally ligated into the
 UniZap XR vector from Stratagene and packaged using
 GigaPack III Gold packaging extracts. Plasmids containing
 cDNA inserts were excised from the recombinant lambda-Zap
 phage using Ex-assist helper phage and propagated in
 XhoI/R cells."

BASE COUNT 181 a 59 c 133 g 199 t

ORIGIN

Query Match 29.1%; Score 177.4; DB 10; Length 572;
 Best Local Similarity 60.3%; Pred. No. 1.1e-42;
 Matches 342; Conservative 0; Mismatches 201; Indels 24; Gaps 2;

QY 26 TATTCAGACATGGAGATGTTTGAATTAATCTCAAGAACGGGAATGATCATATGTTT 85
 DB 555 TTTCATATACACAAAT 496
 QY 86 GAATTCCTGAATTAATCAACATGCGCTGTAATGTCGTATATCCAAATCATGCTCCT 145
 DB 495 GAATTCCTGAATTAATCAACATGCGCTGTAATGTCGTATATATATATATATATATAT 436
 QY 146 GAGCTCTTTTCCACTATTTTCACTCATATACCAAGTCTTCCAGACGATTTCAAGCTC 205
 DB 435 ACATATCTTCCGACATATCTTATTCACAGAAATCAAGTCTTCAACACACTCAAGCTC 376
 QY 206 CTCACATCCCCCAACCAATCCCATTTCAAGTCAAAACATTTAAAGCTTGGCAAGCTC 265
 DB 375 TTGGCTTGTATGAATTCATCTTCATCCAAATCAAAACATTTGAAAGCTTACAAAGATC 316
 QY 266 ACTGCTCTTCTTCTTCAACGCGCGGTAATGACCAATTCCTCTCTCTCTCTCTCTCT 325
 DB 315 CCTCTCAACATCATATCTCATATCTTCACTACCTCTCTCTCTCTCTCTCTCTCTCT 259
 QY 326 ACCATGTGCAAGTGGTTTCGAGATGATTCATTAAGAAACAAGAAATTCATCCAGTTCAA 385
 DB 258 -----CTTTGATATAGATGTGTAAGAAACAAGAAATTCATTAAGATCTAG 214
 QY 386 ACATGATTTTCCCACTAAAGGCTCAATTTCTTCAAGGCTGAATG-----GACGAGACC 439
 DB 213 GCTCTTCTTTTCTACAGAGATTCAGATTCCTTCTTCAAGCTTAATTTGAACTAGATTGCA 154
 QY 440 GATTCCTGAGACCAACATTCAGCTCTCCGACCTAAGCAAGCATCCATCTCTTGTCTC 499

glycine.
1 (bases 1 to 496)
Shoemaker, R., Kelm, P., Vodkin, L., Eppelding, J., Corryell, V., Khanna
A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R./Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Reggen, Invitrogen Corp. 2130
South Memorial Parkway Huntville, AL 35801 For further information
call: (800)-533-4363 or contact via email: cdu@reggen.com
Insert Length: 754 Std Error: 0.00
Seq primer: -40RP from Gibco
High quality sequence stop: 412.
Location/Qualifiers
1. 496
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1015-2732"
/clone_1lb="Gm-c1015"
/tissue_type="Mature flowers, field grown plants"
/lab_host="X110-Gold"
/note="vector: pBluescript II XE, Site 1: EcoRI, Site 2:
XhoI; This cDNA library was constructed from mRNA isolated
from mature flowers of field grown plants. The cDNA
library was prepared using the Stratagene pBluescript II
XR cDNA library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a poly
(dT) sequence with a XhoI restriction site. EcoRI adapters
were ligated to the blunt-ended cDNA fragments followed by
XhoI digestion. The cDNA fragments were directionally
cloned into the EcoRI-XhoI restriction site of the
pBluescript vector. The ligated cDNA fragments were
transformed into X110-Gold host cells. This library was
constructed by Dr. Randy Shoemaker and Dr. John
Eppelding."

BASE COUNT 144 a 88 c 125 g 139 t

ORIGIN

Query Match 23.9%; Score 145.6; DB 10; Length 496;
Best Local Similarity 58.5%; Pred. No. 4.8e-33;
Matches 296; Conservative 0; Mismatches 174; Indels 36; Gaps 1;

54 ACTCAAGAGCGGAATGTATACATCATGTTTGAATTTGAAATCAACATGCGGTCT 113
|||||
476 ATTCAGCTTTGGTAGCAACATCATGCTTGAATCTCTGAAAATCAAGGAGCATATG 417
|||||
114 GAATGGTGTGTAATACCAATCATGCTCTGCACTCTTTTCCATATTTTCACTCAC 173
|||||
416 AATATGGTGTATAGAAAGCAATCATGTTCTGCAATCCTTGGCATGTGTTGATCCAG 357
|||||
174 AATCCAGCTTCCCGACAGATATTCATCACTCTCATCCGCCCAACCATCCCATTC 233
|||||
356 AAGCCAGGCTTTTCAAGACAGATCAAGGCTTGGCTTGATGAAACATCACTCCGTCC 297
|||||
224 AAGTCAACACTTTAAAGCCTTGCAGAGTCACTGTCTTCTTCTTCCACCGCCGCTGA 293
|||||
226 AAGTCAACACTCAAACTCTTGGCAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 262
|||||
294 ATGACCAATCT 353
|||||
261 -----CTCCAACTCACT 213
|||||
354 TCATAAAGACAGAAATTCATCAAGTCAAAACATGTTTCCCACTAAGGGCTCCAT 413
|||||

Db 212 TTTTAAAGACAGAAATTCATCAAGTCAAAACATGTTTCTTTCCACAGATTCAC 153
|||||
Qy 414 TCTTCAGGCTGAATTTGACAGACCCGATTTCTTGAGCAACCAATTCAGTCTCCAGA 473
|||||
Db 152 TCTTATGCTATATTTGGAATTTGCCCTCATCTGAGCATCATTCAGTTCTCAAGG 93
|||||
Qy 474 CTACGAGCCATCTCCATTTCTTGTGAGCTTTTCCGAATAGCGCTTGAATCGGTCTTA 533
|||||
Db 92 CTCACCAATCCGTGCGCCATTCACGTCCTTTTCAAAAATACCTTCAAGTGTGCT 33
|||||
Qy 534 CTACGAGCCATTTTACACCAAA 559
|||||
Db 32 GTTAGAGACATATTTGCAAAATA 7
|||||

RESULT 5
B1425377/c
LOCUS
DEFINITION B1425377.1 Gm-c1076 glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-c1076-2207 5 similar to TR:Q9SVG9 Q9SVG9 CALCIUM-BINDING
PROTEIN-LIKE, mRNA sequence.
ACCESSION B1425377
VERSION B1425377.1 GI:15202391
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 516)
Shoemaker, R., Kelm, P., Vodkin, L., Eppelding, J., Corryell, V., Khanna
A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R./Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Reggen, Invitrogen Corp. 2130
South Memorial Parkway Huntville, AL 35801 For further information
call: (800)-533-4363 or contact via email: cdu@reggen.com
High quality sequence stop: 421.
Location/Qualifiers
1. 516
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1076-2207"
/clone_1lb="Gm-c1076"
/tissue_type="wounded cotyledons"
/dev_stage="11 day old seedlings"
/lab_host="NDH10B"
/note="vector: pBluescript II SK+, Site 1: EcoRI, Site 2:
XhoI; The cDNA library was constructed from mRNA isolated
from 11 day old seedlings treated with that were treated
with 2 ugs/ml of a crude glucan elicitor preparation
isolated from the mycelial walls of Phytophthora sojae.
The library was prepared using the Stratagene pBluescript
II SK(+) library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with an XhoI restriction site. EcoRI
adaptors were ligated to the blunt-ended cDNA fragments
followed by XhoI digestion. The cDNA fragments were
directionally cloned into the EcoRI-XhoI restriction site
of the pBluescript vector. The ligated cDNA fragments were

The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73401

Query Match	23.4%	Score 142.6	DB 10	Length 567
Best Local Similarity	60.6%	Pred. No. 4,1e-33		
Matches 282	Conservative 0	Mismatches 159	Indels 24	Gaps 2

QY	26	TAATCAGACATGGGGATGTTTGAAGATTA	CTCAAGAAAGGAAATGATACATCATGTTTTT	85
DB	453	TTTACATACACAAAATAATATACCTTTAT	CTATACCTTGATATCAACATCATGTGCTT	394
QY	86	GAATTCCTGAAAATTCAAACCATGGCGCT	GTAATTGGTGTGGTAATATACCAATCATGCTCCT	145
DB	393	GAATTCCTGAAAATTCAAAGCCTACCAT	CCAAATTATGATATCAATAAACCGAATCATACTTCT	334
QY	146	GCAAGCTTTTTCACATATTTTCACTCA	CCATACCAGCTTTCGCCAGCAGTAATCAAGCTC	205
DB	333	ACAACTCTTCCCACTACTTTCATCCAA	GAATTCCAAGCTCTTCAACACACACTCAAGCTC	274
QY	206	CTCAATCCCCCACCACCATCCCATTC	CAAGTCAAAACCTTTAAACCTTCGAAAGTGC	265
DB	273	TTGGCTTGTATGATGATTCATCTCCAT	CCAAATCAAAACCTTTGAAATCTTTCACAAATGC	214
QY	266	ACTGTCCTTCTTTCACCCGCGCGTAA	AGAACCAATTCCTCCTCCTCTTCGTGACC	325
DB	213	CCCTCCCAACTCATCATCTTCATCTT	CATCTTCAACCGCTCTATTCGCTTCATTCCTC	157
QY	326	ACCATGTGCCAGTGGGTTGAGATGAT	TATATAAAGAACAAAGATTCATCCAAAGTTCA	385
DB	156	-----CCTTGATATAGGTGTAA	MAAACAAGAAATTCATTGAAGTCTAG	112
QY	386	ACATAGTTTTCCCACTAAGGGCTCA	ATTCCTTCAAGGCTGAATTG-----GACAGACC	439
DB	111	GCCTCTCTTTTCTACAAAGATTC	AGTTCTTTCTTAAGCTAAATTTGTAACTAGATTTGCA	52
QY	440	GATTCCTGGAGCAACCAATTCAGCT	CTCCCAAGCTACAGAGCC	484
DB	51	TATTCCTTGGAGCACAGAGTTTAG	CTCCTCAAGGCTCAAAAAACC	7

RESULT 7	BE568834	667 bp	mRNA	linear	EST 24-MAY-2001
LOCUS	BE568834				
DEFINITION	GM700007B10B5 Gm-r1070 Glycine max cDNA clone Gm-r1070-2529 3', mRNA sequence.				

	ACCESSION	BEE58834			
	VERSION	BEE58834.1			
	KEYWORDS	GI:9984726			
	SOURCE	EST.			
	ORGANISM	Glycine max soybean.			
	REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.			
	AUTHORS	1 (bases 1 to 667) Vodkin,L., Keim,P., Shoemaker,R., Retzel,E., Khanna,A., Coryell,V., Exepding,J., Rapp,C., Shoop,B., Pardinas,J., Liu,L. and Lewin,H. TITLE A Functional Genomics Program for Soybean (NSF 9872565)			
	JOURNAL	Unpublished (1999)			
	COMMENT	Other ESTs: AM156763 corresponding to Gm-cl015-2732 (5') Contact: Vodkin, L.O., PI, A Functional Genomics Program for Soybean (NSF 9872565) Lewin, H. A., Director, Keck Center for Comparative and Functional Genomics.			
	FEATURES	University of Illinois Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA Tel.: (217) 244-6147 Fax: (217) 333-4582 Email: l-vodkin@uiuc.edu This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134. For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact:clones@genomesystems.com or info@genome systems.com web site:www.genomesystems.com Seq primer: 5'-TTTTTTTTTTTTTTT(A/C/G)-3'.			
	SOURCE	Location/Qualifiers 1..667 /organism="Glycine max" /db_xref="taxon:3847" /clone="Gm-r1070-2529" /clone_1fb="Gm-r1070" /note="The library Gm-r1070 is a sequence-driven, rerecked set of 9,216 clones selected from cDNA libraries from various tissues and stages of development of soybean that represent 2,639 sequences from immature cotyledons, 1,770 from immature seed coats, 3,938 from flowers, and 869 from young pods. The 5' ESTs of the source clones form the different libraries was used to select singletons, or a representative of each contig, which were rerecked to form library Gm-r1070. The cDNA clones of the rerecked Gm-r1070 library were then sequenced at the 3' end. The contig analysis to select unique genes was performed by the laboratory of Ernest Retzel, Center for Computational Genomics and Bioinformatics, University of Minnesota, http://www.cbc.um.edu/research/projects/soybean/index.html. Rerecking was performed by Genome Systems, St. Louis, http://www.genomesystems.com, and 3' sequencing by the Keck Center for Comparative and Functional Genomics, University of Illinois, http://www.life.uiuc.edu/blotech/keck.html. Note: The corresponding 5' EST from each clone in the Gm-r1070 library is listed in the 'OTHER EST' field. The detailed information in the source library for each clone can also be obtained by referring to the genome systems clone ID of the original cDNA library that is also listed under 'OTHER EST.'"			
BASE COUNT	201 a	149 C	117 G	176 T	24 others
ORIGIN					
Query Match	23.1%	Score 140.4;	DB 10;	Length 667;	
Best Local Similarity	57.1%;	Pred. No. 2e-31;			
Matches 289;	Conservative	0;	Mismatches 181;	Indels 36;	Gaps 1;
54	ACTCAAGACGGGATGTACATCATCTGTTTGAATTCTTGAAATCAACATGCCGCT	113			
DB	183 ATTCAAGCTTTGGTAGCACAATCATCGTCCTTGTAATTCCTAAAATCAAGGCACATCG	242			
QY	114 GAATTGGTGTGTAATACCAATCATGCTCCGACGCTTTTCCACTAATTTTCACCTCAC	173			

```

Db      243 AAATGTGTCATTAGAAAGAATCATCGTCTCTGAATCCTTGGAATGGTTGTAATCCAG   302
QY     174 ATACCAAGCTTCCCACGACGATTTCAAAGCTCTCACATGCCAACCCATCCCATTC    233
Db      303 AAGCAAGACCCTCTTTCAGACAAGACTCAAGGCTCTGGCTTGATGAAAACCATCTCCGC    362
QY     234 AAGCAAAACACTTAAAAAGCCTTCGAAAGGTCACTGCTTCTTCTTCAACCGCGCGGTGA   293
Db      363 AAGCAAAACACTTAAACGCTTGGCCAAGGTGGT-----              397
QY     294 ATGACCAATTCCTCTCTCTCTCTCTCTCTCTCTCAACCATATGTCAGTAGGTCTGAGATGAT 353
Db      398 -----CTCCAAC TCACTTCCTTTGGCTCTCTCTCTCTCTCTCTCTCTGATTCGATAATTGAG 446
QY     354 TCATAAANAAGACAAAGAATTTCATCCAAAGTTCAAA CAGTGGTTTTCCACTAAAGGCTCCAAT 413
Db      447 TTGTAANAAGAACAAAGAATACTGAAGAACCAAGACTTTCTTTTCCACAAGTGATTTCAAC 506
QY     414 TCTTCAAGGCTGAATTGAGACAGACCCGATTTCTCTGAGACCAACCAATTCACTCTCTCAGA 473
Db      507 TCTCTCTATGCTATATGGAATGCGCCGTCATCTGAAAGCAATGCATTCACGTTCTCTCANNN 566
QY     474 CTAAACGAAGCATCTCATTCATTCTGTGAGGCTTCGAAATP GSGGTTGCAAGTCGGCTCTTA 533
Db      567 NTCAACCAATCGTCGCAATTCAGCTCCACTTTCTCAAAAAATACGCTTCAAGTCACTNGCT 626
QY     534 CTAAAGGGGGACATTTTACCACCAAA 559
Db      627 GNTNMMNMGACACATATTTCGAATA 652

RESULT_8
LOCUS       AM830375/c                     433 bp             mRNA           linear         EST 03-DEC-2001
DEFINITION  sm26a803.y1 Gm-cl028 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: GM-cl028-4997 5' similar to TR:022845 O22845 CALCULUM BINDING PROTEIN ISOLQG.;, mRNA sequence.
VERSION     AM830375
KEYWORDS    soybean.
SOURCE      Glycine max
ORGANISM    Glycine max
REFERENCE   Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
AUTHORS     1.(bases 1 to 433)
            Shoemaker,R., Keim,P., Vodkin,L., Erpellding,J., Corvelli,V., Khanna,A.B., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Stepecoe,M., Theising,B., Allen.M., Bowers,Y., Person,B., Swaller.T., Gibbons.M., Pape,D., Harvey,N., Schurk ,R., Rietter,E., Kohn,S., Shn.T., Jackson.Y., Cardenas,M., McCann ,R., Waterston.R. and Wilson.R.
TITLE       Public Soybean EST Project
JOURNAL     Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public soybean EST project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel.: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
FEATURES             Location/Qualifiers
             source          other_name= "Glycine max"
                             /db_xref="taxon:3847"
                             /clone="GENOME SYSTEMS CLONE ID: Gm-cl028-4997"

```


Db 436 CAGCAGCAGAGAGAGCTCTCGGAGAAATATTAACCATCTCCGTCACATGACATCAATT 377
 Qy 248 AAAAGCCTTCGAGAGTCACTGTCTTCTTCCACCGCCGCCGATGACCAATTCCTC 307
 Db 376 AAAGCCCTCCCAATCGCTTCATCATATCAGCAACAATCATATTTCTTGTAGATCC 317
 Qy 308 CTCCTCTCTTCGTCACCATGTCGAGTGGTTCGAGATGATTCATTAAGAACAA 367
 Db 316 TTGGCTATCTAATAAGC-----GTGCTAGTAAACCGAAG 281
 Qy 368 GAATTCATCAAGTTCAAACATGTTTCCCACTAAGGCTCCCAATTCCTCAAGGCTGA 427
 Db 280 GACTCTCTTAATGAGAGCTCTGTTCACCAAGATCACTTAAGCTATCGGAGTGTG 221
 Qy 428 TTGACAGACCCGATTCCTGAGCAACCAATTCAGCTCCACAGACTAAGACCAATC 487
 Db 220 TTGGCCCAACCGAATTATCAAGATCCAAAGAGCTCGTCGAGGTCACAAAGCCGTC 161
 Qy 488 TCATCTTCTTGAGACTTCTGAAATACCGCTTGCAAGTGGCTTACTTAAGGGGGA 544
 Db 160 TTGGTCTTGTGAGCGCTCTGAAACATTCCTCGAGATCGTTTGTAGTTATGATGA 104

RESULT 12
 A1779002/c 494 bp mRNA linear EST 18-MAY-2001

LOCUS EST5259881 tomato susceptible, Cornell Lycopersicon esculentum cDNA
 clone CLS66H4, mRNA sequence.

ACCESSION A1779002
 VERSION A1779002.1 GI:5277043
 KEYWORDS EST.

ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
 Lycopersicon.

REFERENCE 1 (bases 1 to 494)
 D'Ascenzo, M., He, X., Lyman, J., Matern, A.L., Vision, T., Holt, I.E.,
 Liang, F., Upton, J., Ronning, C.M., Craven, M.B., Fujii, C.Y., Bowman,
 C.L., Nieman, W., Fraser, C.M., Venter, J.C., Tanksey, S.D.,
 Giovannoni, J.J. and Martin, G.B.
 Generation of ESTs from Pseudomonas susceptible tomato
 Unpublished (1999)

TITLE JOURNAL
 COMMENT Clemson University Genomics Institute
 Contact: CUGI
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.

FEATURES
 source location/Qualifiers
 1..494
 /organism="Lycopersicon esculentum"
 /cultivar="R11-13 (Rio Grande x Money Maker)"
 /db_xref="taxon:4081"
 /clone="CLS66H4"
 /clone_1fb="tomato susceptible, Cornell"
 /tissue_type="leaf"
 /dev_stage="4-week old"
 /lab_host="SOLR"
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
 XhoI; cDNA - Tomato Pseudomonas Susceptible EST library.
 Directionally cloned cDNAs inserted into pBluescript SK(-)
) at 5' end with EcoRI and 3' end with XhoI site"

BASE COUNT 175 a 52 c 108 g 159 t
 ORIGIN

Query Match 16.5%; Score 100.6; DB 9; Length 494;
 Best Local Similarity 57.7%; Pred. No. 2,3e-19;
 Matches 206; Conservative 0; Mismatches 139; Indels 12; Gaps 1;
 68 ATGTAACATCATGTTTGAATTTTGAATAATCAACATGCGCTGGAATGGTGTGTA 127

Db 368 ATCAACATCATGCTTTAAACTCCTCATATCAAGTTTCACATCTAAATTTGTGATA 309
 Qy 128 ATACCAATCATGCTCTCGACGTCTTTTCCACTATTTTACCTACATACCAGTCTTC 187
 Db 308 AACATTGATCATACTTTTACATCTTTTCCACATGTTTCATCCACATCTTAATCTTGA 249
 Qy 188 CAGACGATTCAGAGCTCCACATCCGCCAACCCATCCCATTCAGTCAAGCAACACTT 247
 Db 248 CAATGCTCTTGACAGCTCTCAACATATGAAATCCGCTCGTTTAATGAAATACTCT 189
 Qy 248 AAAAGCCTTCGAGAGTCACTGTCTTCTTCTTCCACCGCCGCCGATGACCAATTCCTC 307
 Db 188 AACACTTATAGTATGATCTTTT-----CCAGAAAACATTTTCTCATCTTC 141
 Qy 308 CTCCTCTCTTCTGTCACCATGTCAGGAGTGGTTCGAGATGATTCATTAAGAACAA 367
 Db 140 GTGCTTAATATGCTTACTTATTAATATTTGCTTGTGATGATCATATGAAAGAA 81
 Qy 368 GAATTCATCAAGTTCAAACATGTTTCCCACTAAGGCTCCCAATTTCTCAAGGCT 424
 Db 80 AAATCAATGAGAGCTAGGCTTGTTTTACCAAGTAGCAACTCTAGCTCTCTGTGCT 24

RESULT 13
 BE999195/c 382 bp mRNA linear EST 06-OCT-2000

LOCUS EST430918 GVSN Medicago truncatula cDNA clone pGVSN-15012, mRNA
 sequence.

ACCESSION BE999195
 VERSION BE999195.1 GI:10699471
 KEYWORDS EST.

ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
 Medicago.

REFERENCE 1 (bases 1 to 382)
 Fedorova, M., Pierson, B.L., Samac, D.A., Gantt, J.S., Vance, C.P., Town,
 C.D., Bowman, C.L., Craven, M.B., Cho, J. and Fraser, C.M.
 ESTs from senescent nodules of Medicago truncatula
 Unpublished (2000)

TITLE JOURNAL
 COMMENT Department of Agronomy and Plant Genetics
 Contact: Carroll P. Vance
 University of Minnesota
 411 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108 USA
 Tel: 612 625 5715
 Fax: 651-649-5058
 Email: vance004@maroon.tc.umn.edu
 MTRK4290TK More information is available at:
<http://chryseis.tamu.edu/medicago>
 Seq primer: SKmod (CTA GAA CTA gtc gAT CC).

FEATURES
 source location/Qualifiers
 1..382
 /organism="Medicago truncatula"
 /cultivar="genotype A17"
 /db_xref="taxon:3880"
 /clone="pGVSN-15012"
 /clone_1fb="GVSN"
 /tissue_type="senescent root nodules"
 /dev_stage="mixture of effective nodules from 40 day old
 plants harvested 36 hours post shoot removal and nodules
 collected from 2 month old plants at mid-pod stage"

BASE COUNT 175 a 52 c 108 g 159 t
 ORIGIN
 Query Match 16.5%; Score 100.6; DB 9; Length 494;
 Best Local Similarity 57.7%; Pred. No. 2,3e-19;
 Matches 206; Conservative 0; Mismatches 139; Indels 12; Gaps 1;
 68 ATGTAACATCATGTTTGAATTTTGAATAATCAACATGCGCTGGAATGGTGTGTA 127

packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-ZAP phage using Ex-Assist helper phage and propagated in SOLR cells."

BASE COUNT
ORIGIN

124 a 47 c 87 g 124 t

Query Match 15.6%; Score 94.8; DB 12; Length 382;
Best Local Similarity 58.3%; Pred. No. 1.2e-17;
Matches 218; Conservative 0; Mismatches 132; Indels 24; Gaps 2;

219 AACCATCCCATTTAAGTCAACACCTTTAAAGCTTCGCAAGTCACTGCTCTTCT 278
Db AATCATCTCCATCCAAATCAAAACCTTTGAAAGCTTCACAAAGATCCCTCCCAACTCA 318
219 TCACCGCGCGGTAATGACCAATCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 338
Db TCATCTCATCTTCACTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 273
339 GGGTTCGAGATGATTCATAAAAGAAAGAAATTCATCCAACTTCAAAACATGTTTCCC 398
212 ---TTGATATAGATTTGAAAAGAAAGAAATTCATGAACTGAGCTCTTCTTCT 216
399 ACTAAGGCTCCATTTCTTCAAGGCTGAATG-----GACAGACCCGATTTCTGAGC 452
Db ACAAGAGATTCAGATTTCTTCTGAGCTGAATGATGATGATGATGATGATGATGATG 156
453 AACCAATTCAGCTCCGAGACTAAGAGCATTCATCTTCTGAGCTTCTGAGT 512
Db ACAGATTTAGCTCTCAAGGCTCAAAACCATCAATTTGATCTGATCTTCTCAAAA 96
513 ACGCTTCGAAGTCCGCTCTTCTTCTAAGGGGAGACATTTACCAACCAAGAGATATGATCA 572
Db ATACATGCAATCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 36
573 CCATTAACATCAAA 586
Db CTAAGAGCAATCAAA 22

RESULT 14 BE97593 395 bp mRNA linear EST 06-OCT-2000
LOCUS EST129316 GVSU Medicago truncatula cDNA clone GVSU-1H5, mRNA

DEFINITION BE97593 sequence.

ACCESSION BE97593.1 GI:10697869

KEYWORDS

ORGANISM

Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosida I; Fabales; Fabaceae; Papilionoidae; Trifoliaceae; Medicago.

1 (bases 1 to 395)

Fedorova, M., Pierson, B. L., Samac, D. A., Gant, J. S., Vance, C. P., Town, C. D., Bowman, C. L., Craven, M. B., Cho, J., and Fraser, C. M.

ESTs from senescent nodules of Medicago truncatula

Unpublished (2000)

Contact: Carol P. Vance

Department of Agronomy and Plant Genetics

University of Minnesota

411 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108 USA

Tel: 612 625 5715

Fax: 651-649-5058

Email: Vance004@maroon.tc.umn.edu

University of Minnesota name: M271571e TIGR sequence name: MTRAC397K More information is available at: http://chrysis.tamu.edu/Medicago

Seq primer: SKmod (CTA GAA CTA gtc gat CC).

/db xref="taxon:3880"
/clone="GVSU-1H5"
/clone_1lb="GVSU"
/tissue_type="senescent root nodules"
/dev_stage="mixture of effective nodules from 40 day old plants harvested 36 hours post shoot removal and nodules collected from 2 month old plants at mid-pod stage"
/lab_host="E. coli strain SOLR"
/note="Vector: pBluescript SK +/- Site 1: EcoRI; Site 2: XhoI; cDNA was prepared from polyA+ enriched RNA from the mixture of effective nodules of 40 day old plants harvested 36 hours post shoot removal and nodules collected from 2 month old plants at mid-pod stage. The cDNA was directionally ligated into the Uni-ZAP XR vector from Strategene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-ZAP phage using Ex-Assist helper phage and propagated in SOLR cells."

BASE COUNT
ORIGIN

126 a 53 c 84 g 132 t

Query Match 15.1%; Score 92; DB 12; Length 395;
Best Local Similarity 58.5%; Pred. No. 8.7e-17;
Matches 230; Conservative 0; Mismatches 136; Indels 27; Gaps 3;

219 AACCATCCCATTTAAGTCAACACCTTTAAAGCTTCGCAAGTCACTGCTCTTCT 278
Db AATCATCTCCATCCAAATCAAAACCTTTGAAAGCTTCACAAAGATCCCTCCCAACTCA 324
219 TCACCGCGCGGTAATGACCAATCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 338
Db TCATCTCATCTTCACTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 279
339 GGGTTCGAGATGATTCATAAAAGAAAGAAATTCATCCAACTTCAAAACATGTTTCCC 398
Db ---TTGATATAGATTTGAAAAGAAAGAAATTCATGAACTGAGCTCTTCTTCT 222
278 ---TTGATATAGATTTGAAAAGAAAGAAATTCATGAACTGAGCTCTTCTTCT 222
399 ACTAAGGCTCCATTTCTTCAAGGCTGAATG-----GACAGACCCGATTTCTGAGC 452
Db ACAAGAGATTCAGATTTCTTCTGAGCTGAATGATGATGATGATGATGATGATGATG 279
453 AACCAATTCAGCTCCGAGACTAAGAGCATTCATCTTCTGAGCTTCTGAGT 512
Db ACAGATTTAGCTCTCAAGGCTCAAAACCATCAATTTGATCTGATCTTCTCAAAA 102
513 ACGCTTCGAAGTCCGCTCTTCTTCTAAGGGGAGACATTTACCAACCAAGAGATATGATCA 572
Db ATACATGCAATCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 45
573 CCATTAACATCAAACTTGAAGCTTACCAATT 605
Db CCATTAAGCAATCAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 12

RESULT 15 AM26683 499 bp mRNA linear EST 20-FEB-2001
LOCUS AM266833/c L48-17673 Ice plant Lambda Uni-Zap XR expression library, 48 hours NaCl treatment Mesembryanthemum crystallinum cDNA clone L48-176 5', mRNA sequence.

DEFINITION

AM266833.1 GI:6647268

ACCESSION

KEYWORDS

ORGANISM

Common iceplant.

Mesembryanthemum crystallinum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Caryophyllales; Aizaceae; Mesembryanthemum.

1 (bases 1 to 499)

Cushman, J. C.

An expressed sequence tag database for the common ice plant, Mesembryanthemum crystallinum

Unpublished (1997)

COMMENT

Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu

PCR Primers

FORWARD: T7

BACKWARD: T3

Plate: L48-2 row: G column: 8

Seq primer: T3

High quality sequence stop: 350.

FEATURES

source

Location/Qualifiers

1..499

/organism="Mesembryanthemum crystallinum"

/db_xref="taxon:3544"

/clone_lib="L48-176"

/clone_lib="Ice plant lambda Uni-Zap XR expression library

/tissue_type="leaf, 48 h 0.4M NaCl"

/dev_stage="Six week old"

/note="Vector: lambda Uni-Zap XR, Bluescript SK-; Site_1:

EcoRI; Site_2: XhoI"

Score: 72 c 133 g 156 t

BASE COUNT

138 a 72 c 133 g 156 t

ORIGIN

Query Match

14.4%; Score 88; DB 10; Length 499;

Best Local Similarity 58.9%; Pred. No. 1.6e-15;

Matches 192; Conservative 0; Mismatches 125; Indels 9; Gaps 2;

```
OY 222 CCATCCCATTCATCAAGTCAAGCTTTAAAGCTTCGCAAGTCACTGCTTCTTCA 281
Db 321 CCATACCATTCATCAAGTCAAGCTTTAAAGCTTTATACAGTCACTTCCTCTTC 262
OY 282 CCGCGCGGTGATGACCAATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 341
Db 261 ACAC-----ATTCCATTGATGATCACTACTTGCAATCTGTACTGTCTTCTCC 208
OY 342 TTCGAGATGATTCATTAAGAACAAGATTCATCCAGTTCAAAATGTTTCCACT 401
Db 207 TTGAATATGATTCATAGAGATGAAACTCGTGAATCAAGTCAACTTCACCCACC 148
OY 402 AAGGCTCCATTTCTCAAGGCTGAATTGACAGACCCGATTTCTGGAGCAACCAATTC 461
Db 147 AACGACACAGCTCTCTGCTGAATGACACACACACACACACACACACACATTC 91
OY 462 AGCTCTCGAGCTAAGCAAGCAATTCATCTTGTGAGGCTTCGCAATACGCGTTGC 521
Db 90 AGCTGTGAGGCTCACAAGGCCATCACCATTCTGTGAGGCACTTGAAGATCTATGC 31
OY 522 AAGTGGTCTTACTAAGGGGGGACAT 547
Db 30 AATCTCGGGGTTGAGAGAGACAT 5
```

Search completed: June 29, 2003, 07:22:50
Job time: 1081 secs